2022-1245

United States Court of Appeals for the Federal Circuit

UNILOC 2017 LLC,

Appellant,

– v. –

GOOGLE LLC,

Appellee.

On Appeal from the United States Patent and Trademark Office, Patent Trial and Appeal Board in No. IPR2020-00755

BRIEF FOR APPELLANT

BRETT A. MANGRUM
NATHAN K. CUMMINGS
BRIAN M. KOIDE
RYAN S. LOVELESS
JAMES L. ETHERIDGE
ETHERIDGE LAW GROUP PLLC
2600 East Southlake Boulevard,
Suite 120/324
Southlake, Texas 76092
(817) 470-7249
brett@etheridgelaw.com
nathan@etheridgelaw.com
brian@etheridgelaw.com
ryan@etheridgelaw.com
jim@etheridgelaw.com

Counsel for Appellant

MAY 20, 2022

U.S. Patent 6,366,908, Claim 6 (Appx65)

6. A keyfact-based text retrieving method comprising:

keyfact extracting step for analyzing a document collection and a user query, and extracting keywords without part-of-speech ambiguity from said document collection and said user query, and respectively extracting keyfacts of said document collection and said user query from said keywords;

keyfact indexing step for calculating the frequency of said keyfacts of said document collection and generating a keyfact list of said document collection for a keyfact index structure; and

keyfact retrieving step for receiving said keyfact of said user query and said keyfacts of said document collection and defining a keyfact retrieval model in consideration of weigh factors according to a keyfact pattern and generating a retrieval result.

UNITED STATES COURT OF APPEALS
FOR THE FEDERAL CIRCUIT

CERTIFICATE OF INTEREST

Case Number 2022-1245

Short Case Caption Uniloc 2017 LLC v. Google LLC

Filing Party/Entity Uniloc 2017 LLC

Instructions: Complete each section of the form. In answering items 2 and 3, be specific as to which represented entities the answers apply; lack of specificity may result in non-compliance. **Please enter only one item per box; attach additional pages as needed and check the relevant box**. Counsel must immediately file an amended Certificate of Interest if information changes. Fed. Cir. R. 47.4(b).

I certify the following information and any attached sheets are accurate and complete to the best of my knowledge.

Date: 12/23/2021

Signature: /s/ James L. Etheridge

Name: James L. Etheridge

1. Represented Entities. Fed. Cir. R. 47.4(a)(1).	**2. Real Party in Interest.** Fed. Cir. R. 47.4(a)(2).	**3. Parent Corporations and Stockholders.** Fed. Cir. R. 47.4(a)(3).
Provide the full names of all entities represented by undersigned counsel in this case.	Provide the full names of all real parties in interest for the entities. Do not list the real parties if they are the same as the entities.	Provide the full names of all parent corporations for the entities and all publicly held companies that own 10% or more stock in the entities.
☐ None/Not Applicable	☑ None/Not Applicable	☐ None/Not Applicable
Uniloc 2017 LLC		CF Uniloc Holdings LLC

☐ Additional pages attached

4. Legal Representatives. List all law firms, partners, and associates that (a) appeared for the entities in the originating court or agency or (b) are expected to appear in this court for the entities. Do not include those who have already entered an appearance in this court. Fed. Cir. R. 47.4(a)(4).

☐ None/Not Applicable ☐ Additional pages attached

Jeffrey Huang	Jeffrey A. Stephens	

5. Related Cases. Provide the case titles and numbers of any case known to be pending in this court or any other court or agency that will directly affect or be directly affected by this court's decision in the pending appeal. Do not include the originating case number(s) for this case. Fed. Cir. R. 47.4(a)(5). See also Fed. Cir. R. 47.5(b).

☑ None/Not Applicable ☐ Additional pages attached

Uniloc 2017 LLC v. Google LLC, Appeal No. 21-1509 (Fed. Cir.)		

6. Organizational Victims and Bankruptcy Cases. Provide any information required under Fed. R. App. P. 26.1(b) (organizational victims in criminal cases) and 26.1(c) (bankruptcy case debtors and trustees). Fed. Cir. R. 47.4(a)(6).

☑ None/Not Applicable ☐ Additional pages attached

TABLE OF CONTENTS

Page

TABLE OF AUTHORITIES

Page(s)

Cases:

Statutes & Other Authorities:

STATEMENT OF RELATED CASES

Pursuant to Federal Circuit Rule 47.5(a), Appellant Uniloc 2017 LLC ("Uniloc") certifies that no other appeal from the same proceeding in the United States Patent and Trial Appeal Board ("Board") is or was previously before this Court or any other appellate court. Pursuant to Federal Circuit Rule 47.5(b), the following cases may directly affect or be directly affected by this Court's decision: (1) *Uniloc 2017 LLC et al. v. Google LLC*, Case No. 4:20-cv-05346-YGR (N.D. Cal.) and (2) *Uniloc 2017 LLC v. Google LLC*, Case No. 2-18-cv-00553 (E.D. Tex.).[1]

JURISDICTIONAL STATEMENT

The Board had jurisdiction under 35 U.S.C. § 6 over the petition for *inter partes* review brought by Google LLC ("Google"). The Board issued its Final Written Decision on October 7, 2021 ("Decision"). Appx1-50. Uniloc timely filed its Notice of Appeal on December 8, 2021. Appx524-527. This Court has jurisdiction under 28 U.S.C. § 1295(a)(4)(A) and 35 U.S.C. §§ 319, 141.

STATEMENT OF THE ISSUES

Whether the Board erred, as a matter of law, in adopting a claim construction that failed to recognize dependent claim 7 recites new and distinct acts for the

[1] The litigation filed in the Eastern District of Texas as Case No. 2-18-cv-00553 (E.D. Tex.) was transferred to the Northern District of California as Case No. 4:20-cv-05346-YGR. Appx3605; Appx2. The transferee court dismissed the litigation for lack of subject matter jurisdiction and a notice of appeal was filed on December 31, 2020. *See* Uniloc 2017 LLC v. Google LLC, Appeal No. 21-1509.

overarching "keyfact extracting step" of independent claim 6 that are expressly differentiated from, and hence must be performed in addition to, the differentiated acts recited in claim 6 for the "keyfact extracting step."

Whether the Board erred, as a matter of law, in failing to recognize that the '908 patent provides lexicography for the "keyfact" claim term, which includes a disavowal that expressly limited its scope to the exact form of [object, property].

Whether, as a matter of law, the Board's conclusion of obviousness for all challenged claims is tainted by the Board's application of erroneous claim constructions.

Whether, as a matter of law, the Board erred in failing to recognize patentable distinctions over the cited art arising from a proper understanding of (1) limitations directed to the "keyfact extracting step" of claim 6, including those identified in the Petition as [6a1]-[6a2] and the distinct and additional limitations identified as [7a]-[7d] of claim 7; (2) limitations directed to the "keyfact" term *itself*; and (3) the "keyfact retrieving step" [6c] of claim 6.

STATEMENT OF THE CASE

This is an appeal from the Board's Decision in IPR2020-00755 concluding that claims 6-12 (the "Challenged Claims") of the '908 patent would have been unpatentable as obvious over Braden-Harder in view of the Grossman reference.

A. Overview of the ’908 Patent

The ’908 patent describes novel methods for text retrieval and indexing that is based on what is referred to in the ’908 patent as *keyfacts*. Appx55 (Title), Appx61 1:1-12. The keyfact term is *itself* crucial to understanding the teachings of the ’908 patent. In the context of the ’908 patent, the keyfact term was not a known term of art at the time; rather, it was newly coined by patentee. The ’908 patent states that a “keyfact” is “an important fact contained in sentences which constitute a document,” where the keyfact is “represented by an object and property information through syntactic analysis of a sentence.” Appx61 1:15-18. The ’908 patent underscores and clarifies what is meant by the requisite ordered-pair form of “object and property” by the following disavowal: “[a]ll keyfacts express semantic relation between words in the form of [object, property].” Appx62 4:58-60.

The definitive [object, property] form, universally required for “all” keyfact expression, is a consistent and repeated theme of the ’908 patent. *See*, *e.g.*, Appx61 1:8-10 (“In particular, the method [of the present invention] describe the formalized concept of a document as pair comprising an object that is the head and a property that is the modifier.”); *id.*, 1:7-10 (“The keyfact is represented by an object and property information through syntactic analysis of the sentence.); *see also* Appx63 6:38-44; Table 1 and accompanying description. The requisite [object, property]

form *itself* intrinsically connotes a paired relationship between an "object" that is the head and a "property" that is the modifier. *Id*.

The '908 patent notes various shortcomings of conventional approaches to text storage and retrieval. The '908 patent describes and particularly disparages what it refers to as keyword-based or phrase-based approaches. Appx61 1:19-2:7. As described in the '908 patent, the precision of a keyword-based approach was less than ideal for several reasons. *Id*., 1:19-22. For example, the meaning of the document was not precisely represented. *Id*., 1:22-25. In addition, the keyword-based approach did not precisely reflect intentions behind user queries. *Id*., 1:27-32. This inherent imprecision regarding user queries likewise negatively affected retrieval precision. *Id*., 1:32-34. The '908 patent describes the phrase-based approach as potentially superior to the word-based approach in terms of precision. Nevertheless, the phrase-based approach is described as at least lacking the precision of the disclosed concept-based approach, which can express text by concept units in the form of keyfacts. *Id*., 1:45-49.

The '908 patent teaches a novel approach to text retrieval and indexing that is based on *keyfacts*. *Id*., 1:50-53. Certain embodiments are characterized, in part, as overcoming at least the identified shortcomings of conventional approaches. *Id*. In a keyfact-based text retrieval method, parts of text that represent the same meaning are described as a keyfact in the paired form of [object, property], where similar

concepts are indexed using the same keyfact terms. *Id*., 1:53-55, 60-62. According to the '908 patent, performing indexing and retrieval using keyfacts greatly improves precision and efficiency over conventional approaches, such as the keyword-based and phrase-based approaches disparaged in the background section of the '908 patent. *Id*., 1:55-59.

The Petition purports to challenge claims 6-12 of the '908 patent. Claim 6 is the only challenged independent claim and is reproduced below, with the "keyfact" term central to the dispute on appeal emphasized:

> 6. A keyfact-based text retrieving method comprising:
>
> [6a] keyfact extracting step for [6a1] analyzing a document collection and a user query, and [6a2] extracting keywords without part-of-speech ambiguity from said document collection and said user query, and [6a3] respectively extracting *keyfacts* of said document collection and said user query from said keywords;
>
> [6b] keyfact indexing step for [6b1] calculating the frequency of *said keyfacts* of said document collection and [6b2] generating a keyfact list of said document collection for a keyfact index structure; and
>
> [6c] keyfact retrieving step for [6c1] receiving *said keyfact* of said user query and *said keyfacts* of said document collection and [6c2] defining a keyfact retrieval model in consideration of weigh factors according to a keyfact pattern and [6c3] generating a retrieval result.

Appx65 (emphasis added). The bracketed limitation labels (*e.g.*, "[6a]") are added above for ease of reference herein to the corresponding limitation and are intended to reflect the limitation labeling of claim 6 applied in the briefing below.

B. The IPR Proceeding

Google filed its Petition for *Inter Partes* Review on March 27, 2020, seeking invalidation of claims 6-12 of the '908 patent under eight alleged grounds asserting various references. Appx109. Uniloc filed a Preliminary Response to the petition on July 21, 2020. Appx197. After additional briefing authorized by the Board related to Uniloc's arguments under 35 U.S.C. § 314(a) and the *Fintiv* factors (Appx236, Appx251), on October 7, 2021, the Board granted Google's Petition and instituted *inter partes* review of the Challenged Claims on the following grounds (Appx109):

Claim(s) Challenged	35 U.S.C. §	References
6-12	103	Braden-Harder and Grossman
6-12	103	Braden-Harder, Heidorn, Messerly, and Grossman
7-9	103	Braden-Harder, Grossman, and Kucera
7-9	103	Braden-Harder, Heidorn, Messerly, Grossman, and Kucera
9	103	Braden-Harder, Grossman and Miller
9	103	Braden-Harder, Heidorn, Messerly, Grossman, and Miller
9	103	Braden-Harder, Grossman, Kucera, and Miller
9	103	Braden-Harder, Heidorn, Messerly, Grossman, Kucera, and Miller

Uniloc then filed its Response to the Petition (Appx323), Google filed its Reply (Appx360), and Uniloc filed its Sur-Reply (Appx397). The Board held a virtual hearing by video and telephone on July 15, 2021. Appx466.

The Board issued its Decision on August 9, 2021, determining that all Challenged Claims (6-12) would have been unpatentable as obvious over Braden-

Harder and Grossman. Appx1-50. Accordingly, the Board did not address Google's additional grounds challenging these claims. Appx48 n.22.

After summarizing the filings of record and related matters (Appx2), the '908 patent (Appx3-9), and the instituted grounds of unpatentability (Appx10-11), the Board addressed two claim construction disputes raised in the parties' papers. Appx13-18. First, the Board found it was immaterial whether or not the claim 6 "step for" terms invoke 35 U.S.C. § 112, ¶ 6 ostensibly "because Petitioner has shown that the challenged claims are unpatentable under either construction." Appx17. Second, regarding the disputed "keyfact" term, the Board disagreed with Patent Owner that the specification expressly restricts a "keyfact" to the exact ordered-pair form of [object, property]. Appx18.

The Board recognized that Uniloc had contended "that the combined teachings of Braden-Harder and Grossman do not teach at least the following: (1) the "keyfact" term *itself*; respective acts required for (2) the "keyfact extracting" and (3) the "keyfact retrieving" steps. Appx28 (citing Appx335-353 and Appx409-421).

SUMMARY OF ARGUMENT

The Board erred in adopting a claim construction that failed to recognize dependent claim 7 recites new and distinct acts for the overarching "keyfact extracting step" of independent claim 6 that are expressly differentiated from, and hence must be performed in addition to, the differentiated acts recited in claim 6 for

the "keyfact extracting step." The Board compounded its error in failing to recognize the '908 patent provides lexicography for the "keyfact" claim term, which includes a disavowal expressly limiting scope to the exact form of [object, property].

The Board's conclusion of obviousness for all challenged claims 6-12 is tainted by application of erroneous claim constructions. Misguided by incorrect claim constructions, the Board compounded its error by failing to recognize patentable distinctions over the cited art arising from a proper understanding of (1) limitations directed to the "keyfact extracting step" of claim 6, including those identified in the Petition as [6a1]-[6a2] and the distinct and additional limitations identified as [7a]-[7d]; (2) limitations directed to the "keyfact" term *itself*; and (3) the "keyfact retrieving step" [6c] of claim 6. *See* Appx28 (acknowledging Uniloc had timely raised these disputes in its papers of record) (citations omitted). Accordingly, this Court should reverse the Decision of the Board that Google met its burden to prove claims 6-12 of the '908 patent are unpatentable as obvious.

ARGUMENT

I. Standards of Review

This Court reviews the Board's factual findings for substantial evidence and its legal conclusions *de novo*. *In re Gartside*, 203 F.3d 1305, 1316 (Fed. Cir. 2000). Where there is "no issue . . . as to extrinsic evidence," as is the case here, this Court "review[s] the [Board's] claim construction *de novo*." *In re Cuozzo Speed Techs.*,

LLC, 793 F.3d 1268, 1280 (Fed. Cir. 2015), *aff'd sub nom. Cuozzo Speed Techs., LLC v. Lee*, 579 U.S. 261 (2016) (citing *Teva Pharms. USA, Inc. v. Sandoz, Inc.*, 574 U.S. 318, 331 (2015)).

"Obviousness is a question of law based on underlying findings of fact." *In re Kubin*, 561 F.3d 1351, 1355 (Fed. Cir. 2009). A patent is obvious "if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains." 35 U.S.C. § 103(a) (2012).

II. The Board's Adopted Erroneous Claim Constructions.

The Board's Decision is tainted by its application of erroneous claim constructions, which this Court reviews *de novo* under the circumstances. First, the Board failed to recognize that claim 6 and certain dependent claims each recite distinct *acts* that must all be performed in connection with a corresponding "step for" requirement of claim 6. Second, the Board failed to recognize that the specification contains clear lexicography and a universal disavowal restricting the "keyfact" term to the specific form of [object, property].

A. The Board erred in failing to recognize that the claim language itself expressly differentiates certain acts from one another.

The Board correctly held in its Institution Decision that "35 U.S.C. § 112, ¶ 6 does not apply to claim 6, because the claim contains acts to perform the cited

functions." Appx286; *see also* Appx13, Appx334-339. Google offered no cognizable basis during trial to disturb that holding. The Board ultimately erred, however, in adopting Google's assessment that "the dependent claims do not add new distinct steps or acts, but 'give details of the more general steps previously claimed.'" Appx16 (citations omitted); *cf*. Appx339-343, Appx406-408. This was reversible legal error.

The doctrine of claim differentiation provides that "each claim in a patent is presumptively different in scope." *Trs. of Columbia Univ. in City of New York v. Symantec Corp.*, 811 F.3d 1359, 1370 (Fed. Cir. 2016) (citation omitted). The doctrine is based on "the common sense notion that different words or phrases used in separate claims are presumed to indicate that the claims have different meanings and scope." *Andersen Corp. v. Fiber Composites, LLC*, 474 F.3d 1361, 1369 (Fed. Cir. 2007). Claim differentiation gives rise to a rebuttable presumption for claim construction purposes, especially when comparing differences in scope between an independent claim and its dependent claims: "[T]he presence of a dependent claim that adds a particular limitation gives rise to a presumption that the limitation in question is not present in the independent claim." *Phillips v. AWH Corp.*, 415 F.3d 1303, 1315 (Fed. Cir. 2005).

For ease of reference, the parties have referred to the three distinct acts recited in claim 6 for the "keyfact extract step" as [6a1] ("analyzing a document collection

and a user query"), [6a2] ("extracting keywords without part-of-speech ambiguity from said document collection and said user query"), and [6a3] ("respectively extracting keyfacts of said document collection and said user query from said keywords"). Appx65 9:55-60. Claim 7 adds a new and entirely distinct "analyzing" act as follows: "analyzing morphology of an input sentence and obtaining tag sequences of part-of-speech by attaching part-of-speech tags . . ." (referred to the briefing below, and herein, as limitation [7a]). *Id*., 9:8-11.

Note the express differentiation between the "analyzing" of [6a1] and the "analyzing morphology" of [7a]. In claim 6, the dual-part "analyzing" of [6a1] is directed to both "a document collection and a user query." By contrast, the new and distinct "analyzing" of [7a] is directed to "*morphology* of an input sentence," where [7a] introduces "input sentence" as a new claim element distinct from the "document collection" and "user query" recited in [6a1].

A person of ordinary skill in the art would recognize the "analyzing" of [7a] is not a refinement of the distinct "analyzing" of [6a1] at least because [7a] does not make antecedent reference to the "analyzing" of [6a1] (*e.g.*, by reciting "said analyzing of claim 6"). Rather, claim 7 introduces its differentiated "*analyzing morphology* of an *input sentence*" as a new and distinct act likewise pertaining directly to the overarching "step of keyfact extracting." Under a plain reading of claim 7, therefore, the overarching "keyfact extracting step" requires not only

"analyzing a document collection and a user query" of [6a1], but also additionally requires (among other limitations) performing the *new* and *expressly distinguished* step or act of "analyzing morphology of an input sentence . . ." of [7a].

The Board's error is further underscored by contrasting claims 7 and 8. Claim 8 recites at the outset, "8. The keyfact-based text retrieving method of claim 7, wherein *said step of analyzing morphology* comprises the steps of . . . ." Appx65 10:21-23 (emphasis added). Unlike claim 7, claim 8 makes explicit antecedent reference to the previously recited "analyzing morphology" of claim 7. *Id*. This antecedent reference confirms that the dependent claims make it explicit when their respective limitations simply provide refinements on an antecedent act pertaining to one of the three overarching steps in claim 6. Appx407-408. The absence of such antecedent reference in claim 7 to the first-listed "analyzing" act of claim 6 only underscores the express differentiation between the dual-part "analyzing" of [6a1] and the *new* and *differentiated* "analyzing morphology" of [7a].

The [7a] recitation ". . . obtaining tag sequences of part-of-speech by attaching part-of-speech tags" is likewise expressly differentiated from, and hence cannot properly be conflated with, the distinct [6a2] recitation "extracting keywords without part-of-speech ambiguity from said document collection and said user query." Here, the differentiation is self-evident by comparing, for example, the [7a] limitation of "*obtaining tag sequences*" for an "*input sentence*" with the distinct [6a2] limitation

of "*extracting keywords*" from distinct claim elements differentiated as the "*document collection*" and "*user query*."

Contrary to what the Board suggested, the determinative claim construction issue here is not whether claim 7 "gives details of the more general steps previously claimed." Appx16 (citations omitted). Rather, it is whether claim 7 recites new and distinct acts (for the overarching "keyfact extracting step") that are expressly differentiated from, and hence must be performed in addition to, the acts [6a1], [6a2], and [6a3] recited in claim 6. The Board failed to appreciate that [7a] is not only expressly tethered to the "keyfact extracting step" but also recites new and distinct acts for that step. *See id*. (ironically finding "erroneous" the correct interpretation that claim 7 recites "additional steps or acts" that are *expressly tethered* to the overarching "keyfact extracting step" [6a1]).

The Board evidently failed to grasp the way in which dependent claim 7 fits into the overall method of claim 6. To illustrate the Board's error, the relevant "keyfact extracting step" of claim 6 is reproduced below, with limitation [7a] inserted alongside the other acts of claim 6 ([6a1], [6a2], and [6a3]), which are likewise directed to the same overarching "keyfact extracting step."

> 6. A keyfact-based text retrieving method comprising:
>
> [6a] keyfact extracting step for
>
> [6a1] analyzing a document collection and a user query, and
>
> [6a2] extracting keywords without part-of-speech ambiguity

from said document collection and said user query, and

[6a3] respectively extracting keyfacts of said document collection and said user query from said keywords; [and]

[7a] … *analyzing morphology of an input sentence and obtaining tag sequences of part-of-speech by attaching part-of speech tags . . . .*

Appx65 9:54-60, 10:11-13. The above representation illustrates how [7a] fits into the overarching "keyfact extracting step" as *new* and *additional* limitations, distinct from the other acts [6a1], [6a2], and [6a3] of claim 6. Limitation [7a] of claim 7 clearly requires *both* distinct acts of analyzing, recited in [6a1] and [7a] respectively, where only the latter analyzing (in [7a]) expressly pertains to *morphology* and *tag sequences* for an *input sentence*. In addition, in connection with the new and distinct "analyzing morphology" act, limitation [7a] further requires "obtaining tag sequences …" as an act that is new and distinct from anything recited in claim 6.

The underlying basis for the Board's conclusion of obviousness for all challenged claims 6-12 crumbles if, for example, the limitations expressed in [7a] do not merely *refine* the distinct requirements of [6a1]-[6a3], but also add *new* and *distinct* steps or acts that must be performed for the keyfact extracting step, *in addition to* the three acts [6a1], [6a2], and [6a3] recited in claim 6. *See*, *e.g.*, Appx339-344, Appx406-408, Appx418-419. It likewise follows that the limitations [6a1]-[6a3] each recite respective acts that are required *in addition to* those recited in claim 7; and thus limitations [6a1]-[6a3] must each require something *distinct* and

apart from what is recited in claim 7, in order for limitations [6a1]-[6a3] to each have meaningful effect. *Id*.

When properly construed, therefore, obviousness of acts [6a1], [6a2], and [6a3] cannot be shown in the absence of a theory that recognizes that each one of those acts are differentiated from the limitations of claim 7 (*e.g.*, [7a]), such that [6a1], [6a2], and [6a3] each requires something other than what is *additionally* required in claim 7. The Board's conclusions of obviousness in its Decision are not based on such a theory, nor could it be because no such theory is set forth in the Petition. Rather, as previously noted, the Board based its conclusions of obviousness on the incorrect understanding that "the dependent claims do not add new distinct steps or acts, but 'give details of the more general steps previously claimed.'" Appx16 (citations omitted).

The Board's incorrect claim interpretation constituted reversible legal error that tainted its conclusions of obviousness for all challenged claims 6-12 at issue. As explained further below, the Board compounded its error in overlooking deficiencies of the Petition arising from the proper interpretation that claims 6 and 7 each recite respective, expressly distinguished acts that are each directed to, and individually required for, the overarching "keyfact extracting step."

B. The Board misinterpreted the "keyfact" term.

The Board also erred in failing to recognize that the '908 patent provides

lexicography for the "keyfact" claim term, which includes a disavowal expressly limiting scope to the exact form of "[object, property]" only. Appx18; *see also* Appx35-36 ("We are not persuaded by Patent Owner's arguments that the '908 patent includes a disavowal expressed in the statement 'All keyfacts express semantic relation between words in the form of [object, property].'") (internal quotations omitted)'; *id*. ("Patent Owner directs us to no intrinsic or extrinsic evidence to support its contention that a 'keyfact' must be in the exact paired form '[object, property]' and nothing else.").

It was undisputed below that the keyfact term is coined by the '908 patent, was not a known term of art, and hence required construction by the Board. Appx400 (citing Appx331-332). While Google's Petition purported to interpret "keyfact" to mean "fact contained in sentence(s)," Google applied its alleged interpretation in a manner that would unduly broaden claim scope beyond the lexicography and disavowal in the '908 patent specification.

Uniloc recognized that the obviousness theories of the Petition would effectively requiring expanding claim scope beyond what the intrinsic evidence allows and would risk encompassing what the '908 patent expressly disparages. Appx331-332. In its Response to the Petition, and to focus the dispute, Uniloc proposed that the Board construe "keyfact" to mean "a factual extraction of a sentence which expresses semantic relation between words in the sentence in the

form of [object, property]." Appx334. The Board erred in rejecting Uniloc's construction in favor of interpreting "keyfact" as being wholly unrestricted in terms of its form. *See*, *e.g.*, Appx18, Appx35-36.

That all keyfacts have a requisite form of [object, property] should be deemed admitted by Google. In its claim construction briefing in parallel litigation, Google offered the unequivocal party admission that "the '908 patent teaches that keyfacts are represented 'by a pair comprising an object that is the head and a property that is the modified' in the form of '[object, property].'" Appx953 (citing Appx62 4:58-60 and Appx63 6:38-44); *see also* Appx960 ("keyfact terms are represented in the form of '[object, property]'").

The intrinsic evidence supports interpreting keyfact as being restricted in terms of form, while also precluding the untethered interpretation erroneously applied by the Board. The '908 patent states near the outset that a keyfact is "an important fact contained in sentences which constitute a document," where the keyfact is "represented by an object and property information through syntactic analysis of a sentence." Appx61 1:15-18. The '908 patent clarifies what is meant by restricting "keyfact" to the paired form of "object" immediately followed by "property" by the following additional definitive statement: "[a]ll keyfacts express semantic relation between words in the form of [object, property]." Appx62 4:58-60. This explicit, universal qualification of "[a]ll keyfacts" as having the delimited,

ordered-pair form of "object" immediately followed by "property" is a clear disavowal that limited claim scope. Appx331-334, Appx400-404; *see also X2Y Attenuators, LLC v. Int'l Trade Comm'n*, 757 F.3d 1358, 1362 (Fed. Cir. 2014) (finding a clear disavowal by statement that a feature was "universal" to "all embodiments").

Further underscoring this clear disavowal, the patentee consistently describes the "keyfact" term throughout the '908 patent specification in a manner which always reflects the explicit, universal restriction to the delimited, ordered-pair form of [object, property] only. *See* Appx332-333, Appx401 (collecting citations); *see also See AstraZeneca LP v. Apotex, Inc.*, 633 F.3d 1042, 1052 (Fed. Cir. 2010) ("[W]hen a patentee uses a claim term throughout the entire patent specification, in a manner consistent with only a single meaning, he has defined that term 'by implication.'") (citations omitted). At the outset of the specification, the patentee characterizes the "present invention" as using information described by the "pair comprising an object that is the head and a property that is the modifier." Appx61 1:8-10. A few lines down, the patentee repeats the theme that "[t]he keyfact is represented by an object and property information through syntactic analysis of the sentence." Appx61 1:16-18. With reference to Figure 3, the patentee describes all the keyfacts generated at stage 35 as "hav[ing] the form of [object, property]."

Appx63 6:38-44. The righthand column of Table 1 provides an example list of keyfacts, each one satisfying the requisite form of [object, property]. Appx63.

In describing keyfacts, the '908 patent specification sometimes refers to an instance of the delimited, order-pair form of [object, property] *in its entirety* as a "keyfact term." *See*, *e.g.*, Appx61 2:49-52, Appx63 6:38-40. The specification uses the word "tag" to describe what makes up either the object or property portion of a keyfact term. *See*, *e.g.*, Appx61 2:39-49, Appx62 3:26-36, Appx63 5:61-6:4.

A novel aspect of the claimed invention is that the requisite form of [object, property] *itself* intrinsically conveys certain relational information. Different information is conveyed, for example, when a keyfact *tag* (*e.g.*, KEY1, KEY2, MP, MP1, etc.) is represented as an object as opposed to a property, or when multiple tags (used for either an object or a property) are listed in a particular order. Appx346-347; Appx415-416.

The '908 patent explains this novel aspect achieved by the requisite form [object, property], at least in part, by its description of the examples set forth in Table 1. Appx63 6:13-30. Table 1 separately lists both "[KEY2, KEY1]" and "[KEY2 KEY1, NIL]" as two distinct keyfacts (among two others) collectively representing factual extractions from the statement "the retrieval of information." *Id.*, 6:16-18; *see also* Appx416 (citing Appx63 6:16-18). While the same nonempty tags KEY2 and KEY1 are used in these two example keyfacts, in one instance both tags are

represented as ordered objects ("[KEY2 KEY1, NIL]"), and in the other instance one tag is represented as an object and the other is represented as a property ("[KEY2, KEY1]"). *Id*.

It would be unnecessarily redundant to separately list each one of the above keyfacts (among others) *for the same input sentence* if both keyfacts conveyed the exact same information. That both are included in the same list of keyfacts collectively representing factual extractions of the same input statement confirms that it matters whether a given tag is represented as either an object or a property. As Uniloc explained in its briefing below, and still maintains on appeal, this defining aspect of the restrictive *form* required for all keyfacts accentuates patentable distinctions over the sole obviousness theory the Board considered in its Decision. Appx416 (citing Appx346-347).

The Board compounded its error in failing to recognize additional patentable distinctions arising from the *factual extraction* aspect of a keyfact. Appx333, Appx346-347, Appx409. In contrast to the word-based and phrase-based approaches disparaged in the '908 patent, the keyfact term itself connotes, and the surrounding context set forth in the claim language confirms, that the generation of keyfacts involves *extracting facts*, as opposed to simply repeating verbatim certain words or phrases of an input statement. In the disclosed examples, the factual extractions are represented using *tags* (*e.g.*, KEY1, KEY2, MP, MP1, etc.).

Regarding the generation of representative keyfact tags, the patentee explained that "[p]arts of text that express the same conceptual meaning in the document collection or the query are categorized into the same keyfact type." Appx62 4:61-63. As shown in Table 1, reproduced below, the example tags KEY1, KEY2, MP1, and MP2 each represent a respective derived concept extracted from various input statements.

TABLE 1

Keyfact pattern	Keyfact term list
KEY1 PO KEY2 *(the retrieval of information)*	[KEY2, KEY1], [KEY1, NIL], [KEY2, NIL], [KEY2 KEY1, NIL] [*information, retrieval*], [*information, NIL*], [*retrieval, NIL*], [*information retrieval, NIL*]
KEY1 PO MP KEY2 *(the retrieval of the distributed information)*	[KEY2, KEY1], [KEY1, NIL], [KEY2, NIL], [KEY2 KEY1, NIL], [KEY2, MP]
MP KEY1 PO KEY2 *(the fast retrieval of information)*	[KEY2, KEY1], [KEY1, NIL], [KEY2, NIL], [KEY2 KEY1, NIL], [KEY1, MP]
MP1 KEY1 PO MP2 KEY2 *(the fast retrieval of the distributed information)*	[KEY2, KEY1], [KEY1, NIL], [KEY2, NIL], [KEY2 KEY1, NIL], [KEY1, MP1], [KEY2, MP2]

(Note:
The italic is the examples.)

Appx63 6:14-32.[2]

The '908 patent references Figure 3 in describing a preferred process of using a derived sequence of tags to generate a list of keyfacts, which collectively represent facts extracted from an input sentence. Appx62 4:63-65; Appx63 5:15-18. The patentee explains that subjecting the input statement "the fast retrieval of the distributed information" to "morphological analysis" results in the following sequence: "S (stop-word) A (adjective) NV (vocative noun) PO (possessive preposition) S (stop-word) V-ed (verb) NV (vocative noun). Appx63 5:49-58. The result of morphological analysis is subsequently "converted" into derived *tags* as follows: "NMP KEY PO MP KEY." *Id*. To be clear, the collection of converted tags obtained from the results of morphological analysis is not *itself* a keyfact as claimed.

A "precise sequence of tags is chosen" from the converted keyfact tags, using "part-of-speech tagging." *Id*., 5:61-63. For example, the example processing of the input sentence "the fast retrieval of the distributed information" results in the following final sequence of tags: "MP KEY PO MP KEY." *Id*., 6:1-4. Using this final sequence of tags, "the stage of keyfact pattern extraction 34 searches the keyfact pattern rule 37 and extracts meaningful keyfact patterns." *Id*., 6:5-8. In this example, "[t]he final sequence of tags 'MP KEY PO MP KEY' obtained from 'the

[2] For illustrative purposes, Table 1 of the '908 patent uses italics to indicate certain keywords of a respective input sentence that were used in deriving keyfact tags.

fast retrieval of the distributed information' is applied to the keyfact pattern rule and the keyfact pattern 'MP1 KEY1 PO MP2 KEY2' is the result." *Id*., 6:33-36.

The *keyfact pattern* is not itself a *keyfact* as claimed, but rather it is used in generating keyfacts by application of keyfact generation rules. *Id*., 6:37-55. In the preferred embodiment, for example, generating keyfacts for the input sentence "the fast retrieval of the distributed information" involves outputting a *keyfact pattern* "MP1 KEY1 PO MP2 KEY2," which is applied to *keyfact generation rules* to output the following list of six unique *keyfacts*: "[KEY2, KEY1], [KEY1, NIL], [KEY2, NIL], [KEY2 KEY1, NIL], [KEY1, MP1], [KEY2, MP2]." *Id*.

Note that, in the above example, each one of the six listed keyfacts has the requisite form of [object, property], which is entirely consistent with the lexicography and disavowal of the '908 patent. *Id*. Also note that, in the above example, each one of the six outputted keyfacts provides a respective *factual extraction* of a sentence, at least in that none of the representative internal *tags* (KEY1, KEY2, MP1, MP2, or NIL) repeats verbatim any word or phrase of the input sentence. This is consistent with the patentee's express distinction of the claimed invention over the disparaged word-based and phrase-based approaches in the art. Appx61 1:19-2:7. All other example keyfacts listed in Table 1 consistently apply the same requisite form. Appx63.

The Board erred in failing to appreciate that the above intrinsic evidence

provides lexicography and a clear disavowal that limits "keyfact" to a specific form. Appx35-36. Compounding its error, the Board misplaced the burden with Uniloc to establish that this intrinsic evidence establishes a "keyfact" must be in the exact ordered-pair form of "[object, property]." *See* Appx35 ("Patent Owner directs us to no intrinsic or extrinsic evidence to support its contention that a "keyfact" must be in the exact paired form '[object, property]' and nothing else."). The burden rested with the Petitioner (Google) to defend the unreasonable expansion of claim scope applied in the Petition against the contrary lexicography and disavowal expressed and consistently applied in the '908 patent specification.

The clear lexigraphy and disavowal concerning the keyfact term is not disturbed by the Board's observation that the specification discloses a keyfact may have more than just two internal *tags* (which the Board erroneously referred to as internal *terms*). Appx36. According to the Board, "even the form [KEY2 KEY1, NIL], for example, contains an indicator 'NIL' along with two other terms (KEY2 and KEY1), resulting in a representation that is more than just a pair of terms." *Id*.

The Board failed to appreciate that the example keyfact it had identified ("[KEY2 KEY1, NIL]") applies the requisite *form* of [object, property], where the internal tags used all apply exclusively either to the *object* portion or the *property* portion, and all object tags (KEY2 and KEY1) are separated from, and precede, the property tag (NIL). That the object portion in the cited example consists of two *tags*

does not change the fact that the example still uses the same delimited, ordered-pair *form* of [object, property] and that this requisite form *itself* intrinsically conveys specific semantic relationships. The patentable novelty Uniloc had identified below arises at least in part from the [object, property] form *itself*, regardless of the number of internal tag(s) used for either the object or property portion.

The Board's Decision is tainted by the Board's application of an erroneous construction for the "keyfact" claim term. In finding that the cited art rendered independent claim 6 obvious, the Board applied an overbroad interpretation for the "keyfact" claim term that was untethered to the lexicography and disavowal addressed above. This was reversible error.

III. The Board Incorrectly Determined that claim 6 would have been obvious over Braden-Harder and Grossman.

On appeal, Uniloc challenges the Board's determination that the proposed combination of Braden-Harder and Grossman has been shown to render obvious (1) the recited acts for the "keyfact extracting step" (including those acts recited in claim 6 and the differentiated acts recited in claim 7); (2) limitations directed to the "keyfact" term (when properly construed); and (3) the "keyfact retrieving step." See Appx28 (acknowledging Uniloc had timely raised these disputes in its papers of record) (citations omitted).

A. Level of Ordinary Skill

The Board adopted Google's proposed assessment that a person having

ordinary skill in the art "would have had a Bachelor's degree in computer science, information technology, or the equivalent, and one to two years of experience working with at least one of information retrieval systems, the structure and organization of databases, and natural language processing." Appx12 (citing Appx99). To simplify the issues, Uniloc does not challenge this holding on appeal.

B. The Board erred in failing to recognize deficiencies arising from Google's failure to differentiate between the required acts of claim 6 and the new and distinct acts required by claim 7

As explained above, the Board misinterpreted claim scope in several respects. *See* § II.A, *supra*. Among other reversible misinterpretations, the Board erred in "agree[ing] with Petitioner that the dependent claims do not add new distinct steps or acts, but 'give details of the more general steps previously claimed.'" Appx16 (citing Appx368-372, Appx387-388). In applying its erroneous claim construction, the Board improperly overlooked deficiencies of the Petition arising from Google's improper conflation of expressly differentiated claim limitations.

As noted above, the Board correctly held in its Institution Decision that "35 U.S.C. § 112, ¶ 6 does not apply to claim 6, because the claim contains acts to perform the cited functions." Appx286. As the Board also correctly observed, Uniloc argued in its Response below "that dependent claim 7 recites additional acts required for the 'keyfact extracting step' of claim 6 that are distinct from what claim

6 recites" (as differentiated acts [6a1], [6a2], and [6a3]).[3] Appx15 (citing Appx340-341); *see also* § II.A, *supra*. The Board further summarized Uniloc's Response as contending "that Petitioner erroneously conflates the acts of claim 6 with the acts of the dependent claims 'thereby failing to give meaningful effect to those acts affirmatively recited in claim 6.'" *Id*. (citing Appx341-344).

Google's erroneous conflation is laid bare by the theories set forth in ground 1 of the Petition, which is the sole ground addressed on the merits in the Board's Decision. Appx48, n.22. Misguided by incorrect claim constructions, Google relied in its Petition on the same teaching of Braden-Harder as allegedly satisfying *both* the dual-part "analyzing" of [6a1] ("analyzing a document collection and a user query") *and* the expressly differentiated "analyzing morphology" of [7a] ("analyzing morphology of an input sentence"). *Compare*, *e.g.*, Appx114-115 (alleging Braden-Harder teaches [6a1] by "subject[ing] the documents and user queries to morphological analysis to parse text 'into its constituent words' using 'a stored lexicon'") *with* Appx140 (alleging for the "analyzing" part of [7a] that "Braden-

[3] As previously noted, for ease of reference, the parties referred to the three distinct acts recited in claim 6 for the "keyfact extract step" as [6a1] ("analyzing a document collection and a user query"), [6a2] ("extracting keywords without part-of-speech ambiguity from said document collection and said user query"), and [6a3] ("respectively extracting keyfacts of said document collection and said user query from said keywords"). Appx65 9:55-60.

Harder analyzes morphology of input sentence [sic] to parse text 'into its constituent words' using 'stored lexicon'").

Google likewise relied in its Petition on the same process of Braden-Harder as allegedly satisfying *both* [6a2] ("extracting keywords without part-of-speech ambiguity from said document collection and said user query") *and* the expressly differentiated remainder of [7a] ("obtaining tag sequences of part-of-speech by attaching part-of-speech tags"). *Compare*, *e.g.*, Appx116 (alleging Braden-Harder teaches [6a2] by "applying parts-of-speech tags to each extracted word") *with* Appx140 (alleging for the "obtaining" part of [7a] that Braden-Harder "attaches each parsed word with a part-of-speech tag to obtain tag sequences").

Google's redundant mapping of Braden-Harder onto clearly distinct and differentiated claim limitations reflected Google's incorrect understanding that (1) the "analyzing . . ." of [6a1] somehow encompasses, and requires nothing more than, the differentiated "analyzing morphology . . ." of [7a]; and (2) the "extracting keywords" of [6a2] somehow encompasses, and requires nothing more than, the differentiated "obtaining tag sequences" of [7a]. Such interpretations are fundamentally flawed. *See* § II.A, *supra*.

The Board's Decision did not expressly find any differentiation in the obviousness theory set forth in the Petition concerning acts [6a1], [6a2], and [6a3] with respect to any of the limitations of claim 7 (which are referred to in the Petition

as [7a], [7b], [7c], and [7d]). According to the Board, no differentiation was necessary ostensibly because Uniloc had "fail[ed] to explain sufficiently why we should construe the dependent claims as reciting additional acts to the steps recited in claim 6." Appx16. The Board appeared to have misplaced the applicable burden here, which rested with Google to at least explain why the presumption arising from claim differentiation is rebutted. *See* § II.A, *supra*.

The record simply cannot support a finding of obviousness in the absence of any recognized theory of record that correctly interpreted limitations [6a1], [6a2], [6a3], and [7a] as each reciting *distinct*, *differentiated* acts that are all collectively directed to the same overarching "keyfact extracting step." The Board's Decision does not offer any conclusion of obviousness allegedly based on such an interpretation, nor could it have properly done so because the Google's Petition failed to advance any obviousness theory under such an interpretation. Because the Board's erroneous conclusions of obviousness as to claims 6 and 7 are tainted by incorrect claim constructions, and given all claims challenged in the Petition depend (directly or indirectly) from claim 6, the error identified herein warrants reversing the Board's Decision in its entirety.

C. The Board erred in failing to recognize deficiencies arising from a proper understanding of the "keyfact" claim term

The Board erred in failing to recognize patentable distinctions between the *ordered-pair* form of [object, property] strictly required for all "keyfact" terms and

Braden-Harder's distinguishable *three-part* "triples" form relied upon in the Petition. *See*, *e.g.*, Appx17-18 (expressing the Board's final construction of the "keyfact" term); § II.B, *supra* (addressing error in the Board's construction); Appx345-347, Appx410-417 (addressing deficiencies in the Petition arising from a proper construction). The Board erroneously found that "Braden-Harder's triples are in the form [object, property] because they are in the form, for example, [object, Mods, modifier], where each triple contains two node words linked by a semantic relationship." Appx34-35 (citation omitted). The Board acknowledged this finding is tethered to, and rests upon, the Board's claim interpretation that the "keyfact" claim term is not strictly limited to the exact ordered-pair form of [object, property]. Appx35. The Board's obviousness conclusion as to claim 6 is therefore tainted by the application of an erroneous claim construction. *See* § II.B, *supra.*

As explained previously, the '908 patent defines and strictly limits all keyfacts to the ordered-pair form of [object, property]. *See* § II.B, *supra*. In characterizing the "present invention," the '908 patent states that the keyfact form *itself* intrinsically conveys *factual concepts* "as *a pair* comprising an object that is the head and a property that is the modifier." Appx61 1:6-12 (emphasis added). This *paired* or *two-part* form is plainly distinguishable on its face from the *three-part* "triples" form of Braden-Harder relied upon in the Petition, which links together two words of a statement (copied verbatim) by interposing therebetween *a labeled relation* as a

requisite *third* part of the "triples" form.

Scrutiny of the Braden-Harder disclosure only underscores the distinction. Near the outset of the disclosure, Braden-Harder describes its "triples" form as requiring the "form of 'word-relation-word'" in order to operate as intended. Appx1442 5:16-25. More specifically, the "triples" form applied in Braden-Harder expressly requires three parts: (1) a first word and (2) a middle relation label which by its content interlinks the first word with (3) a second word. *Id*.; *see also* Appx1446 14:3-4 ("Each triple contains two node words as depicted in the graph linked by a semantic relationship therebetween."). Braden-Harder's middle *relation label* is essential in that it defines how the first and second words are related. *Id*.

By way of contrast, an example patentable feature of a keyfact, as disclosed and claimed in the '908 patent, is that its requisite paired form of [object, property] *itself* intrinsically expresses semantic relation between the paired portions. Among other technical advantages, this form obviates the need, for example, to expressly label a semantic relation through inclusion of *third* portion interposed between the object and property. *See* Appx413; *see also* Appx435-347. In Braden-Harder it is not the *form* of the "triples" that conveys semantic relationships between words, much less a paired form. Rather, each given Braden-Harder three-part "triple" expresses the relation by the particular *content* of the interposed *relation label* sandwiched between the two words that are linked.

Another example patentable feature of a keyfact is its intrinsic expression of *concepts* in terms of *factual extractions*. Appx333, Appx346-347, Appx409; *see also* § II.B, *supra*. As noted above, Braden-Harder discloses that each of its "triples" repeat verbatim two actual words of a given statement, together with a relation label interposed therebetween. This is more analogous to the word-based and phrase-based approaches disparaged in the '908 patent. *See* Appx61 1:19-45.

By contrast, the extraction of keyfacts, as disclosed and claimed, necessarily involves extracting *keyfacts* from *keywords*—i.e., the key*facts* and key*words* are not one and the same. *See* § II.B, *supra*. Consistent with the specification, this distinction is made explicit in claim 6 at least in the following recitation: "extracting keywords . . . from said document collection and said user query, and respectively *extracting keyfacts* of said document collection and said user query *from said keywords*." Appx65 9:57-60. The distinction is likewise expressed, for example, in claim 7—*e.g.*, by the additional recited acts directed to "selecting a tag sequence," "extracting a keyfact pattern by applying said tag sequence," and "applying said keyfact pattern to a keyfact pattern generation rule and generating a keyfact list." *Id*., 10:11-19.

The Board also erred in finding "it would have been obvious to generate other logic forms that would not include relations to enable structural paraphrasing." Appx38; *cf*. Appx413-414. In its Petition, Google did not expressly advance the

argument that it would have been obvious to modify Braden-Harder to completely exclude altogether the middle relation portion in its entirety from its "triples" form, thereby converting the self-defining "triple" form to *a different form* having less than three parts. At best, in citing to a portion of Braden-Harder that still relied on and applied the same "triple" form, the Petition ambiguously suggested it would have been obvious to not include any *labeled* relation (*e.g.*, presumably by leaving the label blank). Appx124 (citing Appx1452 25:49-60). As discussed further below, this clearly is not what the cited portion of Braden-Harder discloses.

But even if it would have been obvious to modify Braden-Harder's "triples" to omit all relation labels, which remains contested, this still would not address certain distinctions of record. Such a modification would still not satisfy the required form of [object, property], for the same reasons Uniloc consistently maintained throughout trial. In addition, such an undisclosed modification would not affect distinctions arising from the requirement that keyfacts intrinsically express concepts in terms of factual extractions, as opposed to simply repeating verbatim the actual words of a given statement.

The Board also failed to appreciate that such a modification would only further distinguish Braden-Harder from claim language directed to the "keyfact" term. As explained above, the "keyfact" term is properly construed as expressing *semantic relations*. *See* § II.B, *supra*. This clearly cannot be shown obvious by a

modification that would omit the relation label from the three-part Braden-Harder "triple"—i.e., precisely what Braden-Harder expressly relied upon to portray a semantic relationship between the two words serving as bookends of the "triple."

Google also failed to defend such a modification of Braden-Harder as not rendering the disclosure inoperable for its intended purpose. Appx413-414 (citing *In re Gordon*, 733 F.2d 900, 902 (Fed. Cir. 1984) (if a proposed modification would render the prior art invention being modified unsatisfactory for its intended purpose, then there is no suggestion or motivation to make the proposed modification). Even in the broadest sense disclosed, Braden-Harder defines its "logical form" as "words . . . *linked by labeled relations*." Appx1445 11:42-44 (emphasis and underlining added). Throughout the disclosure, Braden-Harder repeatedly and consistently emphasizes the importance of the relation label, which makes up the middle portion of the requisite "triples" form and which is interposed between the two words linked by that relation label.[4] Indeed, Braden-Harder relies on the

[4] *See*, *e.g.*, Appx1424 Abstract ("Each such logical form encodes, in a word-relation-word manner, semantic relationships, particularly argument and adjunct structure, between words in a phrase."); Appx1445 11:41-43 ("Broadly speaking, a logical form is a directed acyclic graph in which words representing text of any arbitrary size are linked by labeled relations."); Appx1446 13:29-31("To identify all the semantic relationships in an input string, each node in the syntactic parse tree for that string is examined."); *Id.*, 14:15-17 (describing "three logical form constructions for which additional natural language processing is required to correctly yield all the logical form triples"); Appx1451 23:29-31 ("all remaining document records are assigned a score as defined above and based on the relation type(s) of matching triples").

interposed, labeled relation portion of each "triple" not only to express the relation between the two words on either side of the "triple," but also to determine a "predefined weight." Appx1442 5:41-44; *see also* Appx1447 16:19-21.

Braden-Harder further discloses a preferred match process that relies on *all three* portions of the "triple" form, including a match of the relation label *itself*: "Unless all three elements of one triple identically match corresponding elements of another triple, these two triples do not match." Appx1447 16:13-15. This disclosure directly contradicts Google's argument that a person of ordinary skill would have recognized the removal of Barden-Harder's middle relation label altogether would have been preferable. Appx124.

In its Decision, the Board pointed to the Braden-Harder disclosure that "the criteria for determining a match . . . can be relaxed to compass paraphrases as matching." Appx39 (citing Appx1452 25:50-53). The Board appeared to overlook, however, that Braden-Harder also discloses (in the same passage) that Braden-Harder describes the "relaxed" criteria as still applying to the *three-part "triples"* form and still involving matching the *semantic relation* of similar words. Appx1452 25:49-65. In other words, even when optionally broadening the criteria to consider similar *linked words* of the "triples" under comparison, the matching still requires comparing respective relation labels.

Google also failed to reconcile its modification theory, which would have

allegedly omitted a relation label, with how Google addressed the remainder of claim 6. As the Board observed, in addressing the "keyfact retrieving step," Google relied on Braden-Harder's alleged teaching of "generating logical form triples from the user query block and receiving logical form triples"—i.e., the same three-part "triples" form necessarily including a middle relation label which interlinks first and second words together. Appx42 (citations omitted). The Board further observed that Google relied upon predefined weights based on the particular *relation label* used for a given triple. Appx42-43 (citations omitted). Absent from the record is any explanation, or any express finding from the Board, reconciling how it would have been obvious, or even possible, to *omit* the crucial relation labels from the "triples" form while at the same time still *rely* upon the relation labels for establishing predefined weights.

Thus, Braden-Harder speaks for itself in emphasizing that the non-empty, relation label used in the middle portion of all three-part "triples" is a necessary feature of the disclosure. The record simply fails to establish that omission of the middle relation label from the three-part form of all Braden-Harder's "triples," in a manner not disclosed by that reference or by any other cited art, would not have rendered Braden-Harder unsatisfactory for its intended purpose.

D. The Board erred in placing the burden on Uniloc to defend an erroneous construction Google had applied in its Petition

The Board erred in finding Uniloc failed to persuasively defend an aspect of

the construction *Google had advanced in its Petition* for the "keyfact extracting step." In its Decision, the Board held "we do not agree that the 'weigh factors' must be determined experimentally or cannot be predetermined weights as Patent Owner asserts." Appx44. The Board failed to appreciate it was Google who had advanced the argument that claim 6 allegedly requires *determining keyfact weight constants* as one of several acts corresponding to the "keyfact extracting step" of claim 6, under the step-plus-function construction applied in the Petition. While Uniloc had contested Google's assertion that the claim 6 invoked 35 U.S.C. § 112, ¶ 6, Uniloc nevertheless identified fatal deficiencies arising under Google's step-plus-function construction. This did not somehow shift the burden to Uniloc to defend any aspect of Google's step-plus-function construction *itself*.

Google's Petition applied a step-plus-function construction for the "keyfact retrieving step" that identified the following corresponding requisite acts as follows: "[t]he specific acts and structures to achieve these claimed functions are disclosed in the specification as shown in Figs. 5-6 and at 7:35-8:24, including the keyfact retrieving device 13 (see Figs. 1-2) that forms a document and user query vector 54 (blue), *determines keyfact weight constants 55 (orange)*, calculates a document query weight 56 (yellow), and ranks retrieved documents that are displayed 57 (green) using a keyfact retrieval model (Fig. 6)." Appx107 (emphasis altered). Without conceding as to the correctness of Google's construction, Uniloc had simply

argued that the alleged requisite act (among others) of *newly determining keyfact weight constants for a given document and user* query, as disclosed in the '908 patent, is plainly distinguishable from the cited disclosure in Braden-Harder of assigning *predefined* weights for the collection of relation labels used by the "triples" form. Appx351-352; Appx419-421.

Uniloc further explained why the patentee's description of the cited act in question (determining keyfact weight constants) is distinguishable on its face from the cited disclosure of Braden-Harder relied on in the Petition. *Id*. Consistent with the alleged requirement of *newly determining keyfact weight constants*, Uniloc had argued, for example, that the '908 patent teaches its "keyfact weight constants" are not *predefined* but rather they are *newly determined* "experimentally" based on the present "distribution of keyfact patterns" at issue. *Id*. (citing Appx64 7:45-67).

Uniloc also explained why the corresponding disclosure in the '908 patent was plainly distinguishable from the cited disclosure of Braden-Harder. *Id*. The Petition pointed to Braden-Harder's Figure 8A and Tables 2 and 3, which discloses assigning *predefined* weights to all *relationship labels* used for the "triples" form. Appx352 (citing Appx135, which cites Appx1433 at Fig. 8A and Appx1446 at Tables 2-3). As shown Braden-Harder's Figure 8A, reproduced below, the relationship labels "Dobj," "Dsub," "Ops," and "Nodj" are assigned the *predefined* weights of "100," "75," "10," and "10," respectively.

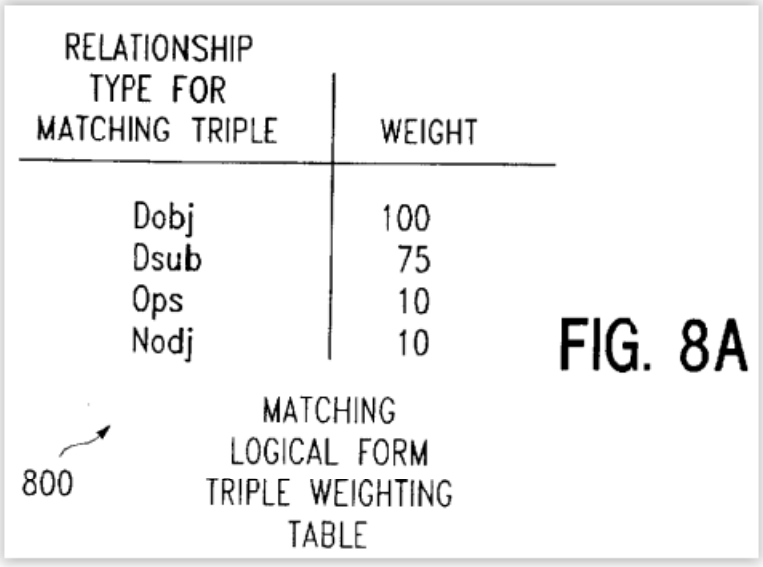


FIG. 8A

Appx1433 at Fig. 8A. Braden-Harder's Figure 5B, reproduced below, illustrate examples of "logical form triples" that each include one of the above relation labels interposed between two words.

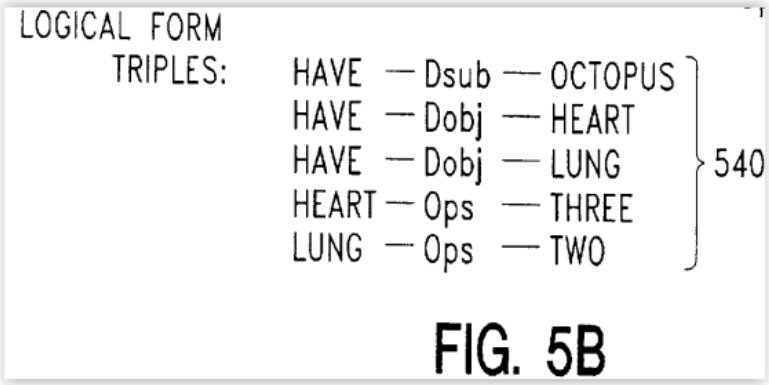


FIG. 5B

Appx1433 at Fig. 5B.

Uniloc explained in its briefing below why simply assigning a respective *predefined* weight, for each relation label used by Braden-Harder's "triple" form, has not been shown to render obvious at least the "determining" act Google had identified in its step-plus-function construction—i.e., the act directed to *newly*

determining keyfact weight contestants "experimentally" based on the present "distribution of keyfact patterns" at issue (i.e., on more than one keyfact pattern alone). *Id*. (citing Appx64 7:45-67). In raising this rebuttal to the theory expressed in the Petition, the burden clearly did not shift to Uniloc to defend any aspect of the step-plus-function construction *advanced by Google*. The Board erred in either misapprehending or overlooking this patentable distinction; it further erred in failing to appreciate the distinction that arises from a "determining" aspect of the step-plus-function construction *applied in the Petition*.

CONCLUSION

For the foregoing reasons, the Court should reverse the Board's Decision that claims 6-12 of the '908 patent are unpatentable as obvious.

Respectfully submitted,

/s/ Brett A. Mangrum
Brett A. Mangrum
Nathan K. Cummings
Brian M. Koide
Ryan Loveless
Jim Etheridge
ETHERIDGE LAW GROUP
2600 East Southlake Boulevard
Suite 120-324
Southlake, TX 76092
P. 817-470-7249
brett@etheridgelaw.com
nathan@etheridgelaw.com
brian@etheridgelaw.com
ryan@etheridgelaw.com
jim@etheridgelaw.com

Attorneys for Appellant-Patent Owner Uniloc 2017 LLC

ADDENDUM

ADDENDUM TABLE OF CONTENTS

Trials@uspto.gov
571-272-7822

Paper No. 28
Date: October 7, 2021

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

GOOGLE LLC,
Petitioner,

v.

UNILOC 2017 LLC,
Patent Owner.

IPR2020-00755
Patent 6,366,908 B1

Before SALLY C. MEDLEY, KRISTEN L. DROESCH, and SHEILA F. McSHANE, *Administrative Patent Judges*.

MEDLEY, *Administrative Patent Judge*.

JUDGMENT
Final Written Decision
Determining All Challenged Claims Unpatentable
35 U.S.C. § 318(a)

I. INTRODUCTION

Google LLC ("Petitioner") filed a Petition for *inter partes* review of claims 6–12 of U.S. Patent No. 6,366,908 B1 (Ex. 1001, "the '908 patent"). Paper 1 ("Pet."). Uniloc 2017 LLC ("Patent Owner") filed a Preliminary Response. Paper 8 ("Prelim. Resp."). In accordance with Board authorization, Petitioner filed a Reply to the Preliminary Response (Paper 9) and Patent Owner filed a Sur-reply (Paper 11). Upon review of the papers, we instituted *inter partes* review, pursuant to 35 U.S.C. § 314, as to claims 6–12 based on the challenges set forth in the Petition. Paper 14 ("Decision to Institute" or "Dec.").

Subsequent to institution, Patent Owner filed a Patent Owner Response (Paper 17, "PO Resp."), Petitioner filed a Reply to Patent Owner's Response (Paper 19, "Pet. Reply"), and Patent Owner filed a Sur-reply (Paper 20, "Sur-reply"). On July 15, 2021, we held an oral hearing. A transcript of the hearing is of record. Paper 27 ("Tr.").

For the reasons that follow, we conclude that Petitioner has proven by a preponderance of the evidence that claims 6–12 of the '908 patent are unpatentable.

A. Related Matters

Petitioner and Patent Owner indicate that the '908 patent is the subject of court proceeding, *Uniloc 2017 LLC v. Google LLC*, No. 2:18-cv-00553 (E.D. Tex.). Pet. 93; Prelim. Resp. 3. Petitioner's motion to transfer the district court case to the Northern District of California was granted. Ex. 1046. According to Petitioner, "the Northern District of California found that at least Fortress Credit Co. LLC held sufficient rights in the challenged patent that Uniloc lacked standing to sue. The court then

dismissed the litigation for lack of subject matter jurisdiction." Paper 18, 1 (citing *Uniloc 2017 LLC v. Google LLC*, No. 4:20-cv-05345-YGR (N.D. Cal. Entered Dec. 22, 2020 (single order addressing eleven cases)). Petitioner states "Uniloc filed a notice of appeal on December 31, 2020." *Id.*

B. The '908 Patent[1]

The '908 patent describes a keyfact-based text retrieval method and a keyfact-based text index method. Ex. 1001, 1:6–7. The '908 patent states that a keyfact is "an important fact contained in sentences which constitute a document," where the keyfact is "represented by an object and property information through syntactic analysis of a sentence." *Id.* at 1:15–18. As described in the '908 patent, a keyword-based text retrieval method was the mainstream in conventional text retrieval methods, but the precision of a keyword-based text retrieval method was less than ideal for several reasons. *Id.* at 1:19–22. First, the meaning of the document was not precisely represented because the document was represented by keywords, which are nouns. *Id.* at 1:22–25. Second, when a query included a natural language phrase, sentence, or keywords, the intention of the user's query was not reflected precisely in a keyword-based text retrieval method. *Id.* at 1:27–32. Therefore, the keyword-based text retrieval method had a fundamental limitation in retrieval precision because it performed document retrieval by keywords. *Id.* at 1:32–34. Phrase-based text retrieval methods perform more precise text retrieval than the keyword-based text retrieval method, but performs less precise text retrieval than a concept-based text retrieval method, which expresses text by concept units. *Id.* at 1:45–49.

[1] The '908 patent appears to be expired.

The '908 patent further describes a new approach to keyfact-based text retrieval that overcomes the shortcomings of the keyword-based text retrieval method and generalized phrase-based text retrieval methods. *Id.* at 1:50–53. In a keyfact-based text retrieval method, parts of text that represent the same meaning are described as a keyfact, where the phrases or words having the same meaning are indexed as the same indexing terms. *Id.* at 1:53–55, 1:60–62. According to the '908 patent, since the keyfact-based retrieval method is a concept-based retrieval method, indexing and retrieval of the keyfact-based retrieval method are performed with the unit of the keyfact, and precision of the retrieval is greatly improved. *Id.* at 1:55–59.

A block diagram of a keyfact-based text retrieval system is illustrated in Figure 1, reproduced below.

FIG. 1

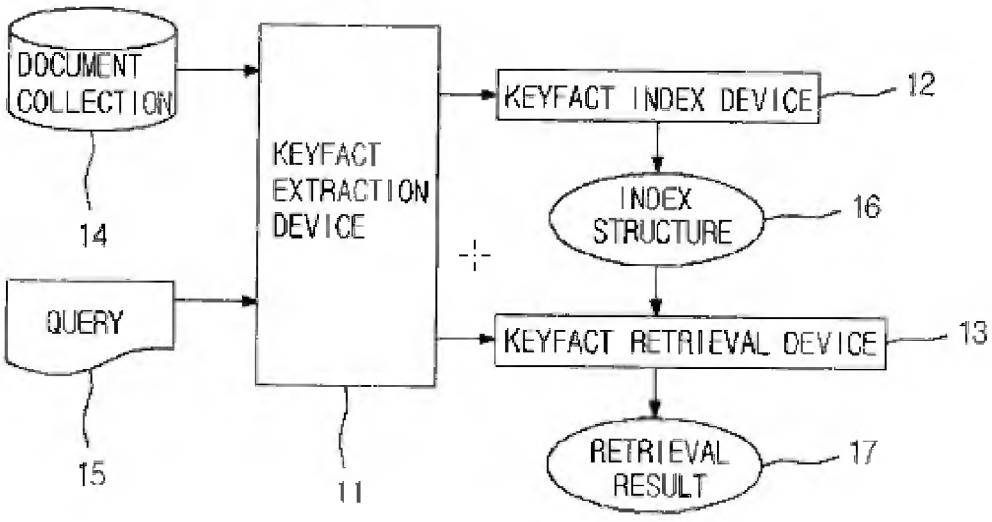


Figure 1 is a block diagram illustrating a keyfact-based text retrieval system. *Id.* at 4:22–23. The keyfact-based text retrieval system comprises

keyfact extraction device 11, keyfact index device 12, and keyfact retrieval device 13. *Id.* at 4:23–26. The '908 patent describes that once a document collection 14 or a query 15 is given, the keyfact extraction device 11 extracts words without ambiguity by performing morphological analysis and tagging. *Id.* at 4:42–45. The keyfact generation rule is applied to the words and then the keyfacts are extracted. *Id.* at 4:45–46. The keyfact index device 12 indexes the document collection 14 or the query 15 with the unit of keyfact and calculates the frequencies of the keyfacts. *Id.* at 4:47–49. The frequencies of the keyfacts are stored into the index structure 16 with the document ID information. *Id.* at 4:49–51. The keyfact retrieval device 13 orders documents using a similarity calculation method and shows retrieval results. *Id.* at 4:51–53.

A block diagram of a keyfact extraction device of a keyfact-based text retrieval system is illustrated in Figure 3, reproduced below.

FIG. 3

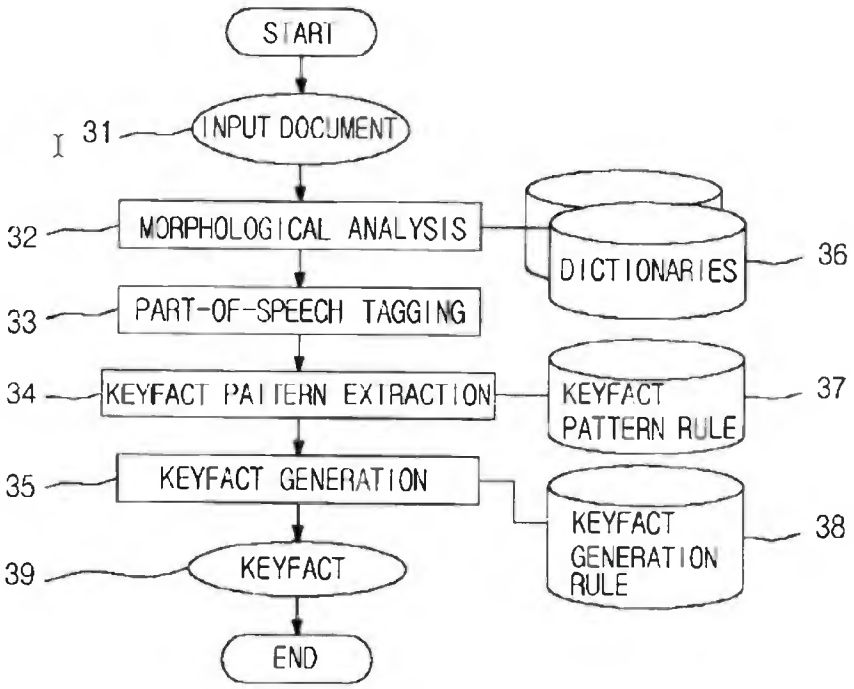


Figure 3 illustrates a block diagram of keyfact extraction device 11 that analyzes a document and generates keyfacts through the processes of morphological analysis, part-of-speech tagging, keyfact pattern extraction, and keyfact generation. *Id.* at 5:15–18. The '908 patent describes that a document is supplied at stage 31 and morphological analysis is performed at stage 32. *Id.* at 5:19–20. More specifically, a sentence in the document is divided into words, the morphological analysis is performed with dictionaries 36 at stage 32, and various results are obtained. *Id.* at 5:20–22, 5:59–60. At stage 33, part-of-speech tagging is performed, where a precise sequence of tags is chosen among the various results of the morphological analysis. *Id.* at 5:61–63. Once the final sequence of tags is obtained, the

stage of keyfact pattern extraction 34 searches the keyfact pattern rule 37 and extracts meaningful keyfact patterns necessary for keyfact generation. *Id.* at 6:5–8. "Keyfact terms that have forms of [object, property] are generated as to the input keyfact pattern at the stage of the keyfact generation 35 by searching the keyfact generation rule 38," resulting in a keyfact list 39. *Id.* at 6:38–41, 51.

A block diagram of a keyfact index device of a keyfact-based text retrieval system is illustrated in Figure 4, reproduced below.

FIG. 4

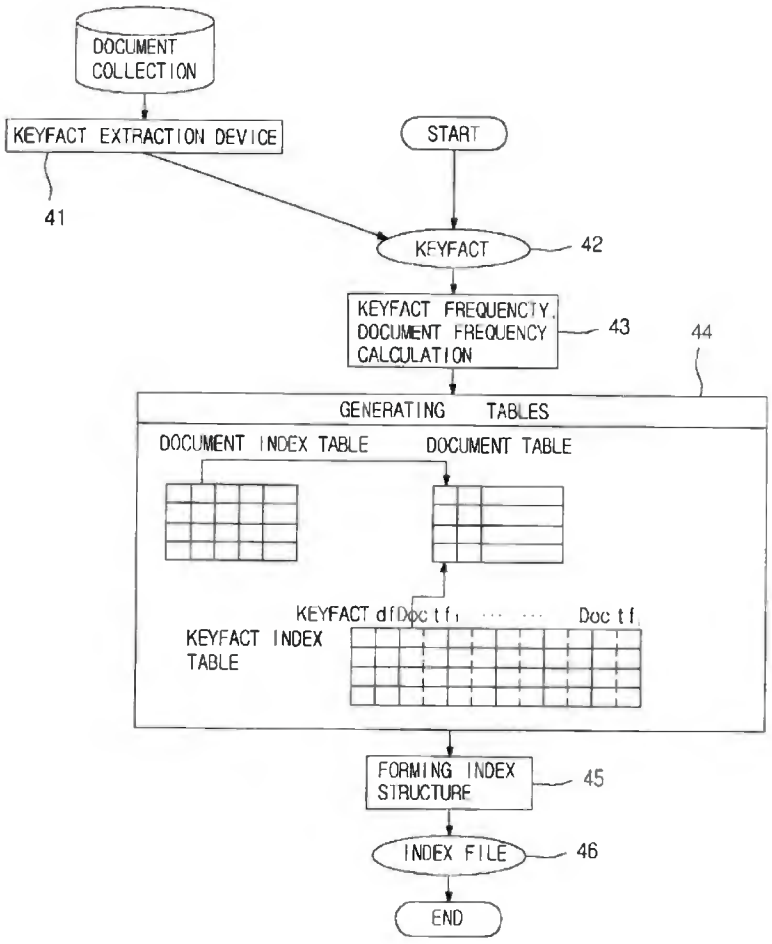


Figure 4 illustrates a block diagram of keyfact index device 12 that calculates statistical frequencies of keyfacts in a document obtained from the keyfact extraction device 11, and that forms an index structure. *Id.* at 6:58–60. The '908 patent describes that, for each document, a keyfact frequency and document frequency of a keyfact 42 are calculated at 43 in order to obtain the frequency information of the keyfacts. *Id.* at 6:65–67. Next, supplementary tables, such as a document index table, a document table, and a keyfact index table, are generated to form an efficient index structure 44. *Id.* at 7:1–3. Subsequently, an index structure is formed in a unit of the keyfact at 45, and an index file is stored at 46. *Id.* at 7:10–11.

A block diagram of keyfact retrieval device of a keyfact-based text retrieval system is illustrated in Figure 5, reproduced below.

FIG. 5

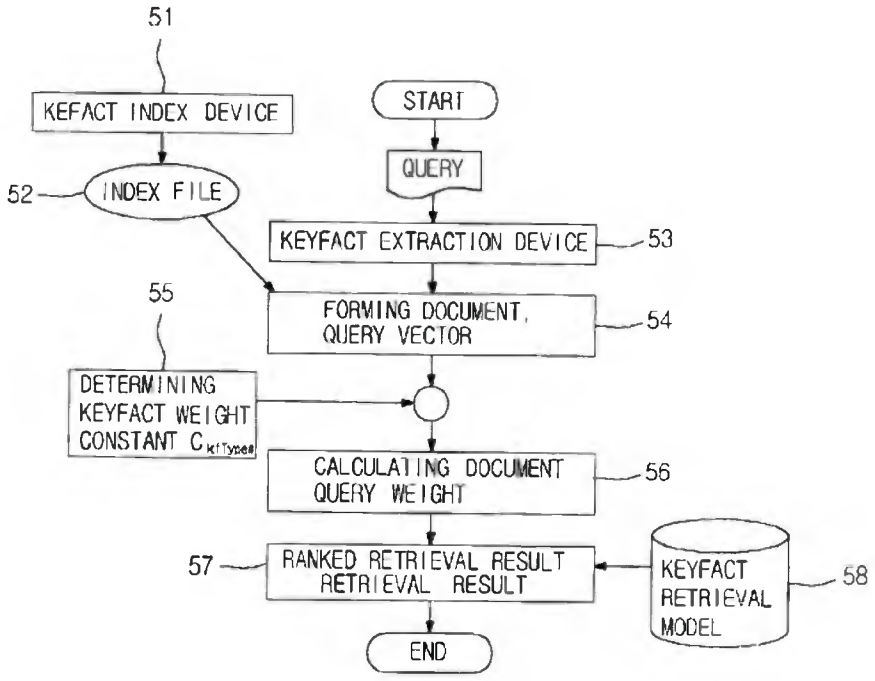


Figure 5 illustrates a block diagram of keyfact retrieval device 13 that, at 54, forms a document vector and query vector with a keyfact supplied from a keyfact extraction device 53 and an index file 52 generated by a keyfact index device 51. *Id.* at 7:35–39. At 55, the keyfact weight constants ($C_{KfType\#}$), which are fit for the attribute of a document collection, are determined. *Id.* at 7:40–41. Subsequently, at 56, the keyfact weights are calculated from the document and query vector. *Id.* at 7:42–43. Further, at 57, the similarity of each document appropriate for the query is calculated by employing a keyfact retrieval model 58, where the result of a similarity calculation determines an order of appropriate documents. *Id.* at 8:20–24.

C. Illustrative Claim

Petitioner challenges claims 6–12 of the '908 patent. Pet. 1. Claim 6 is an independent claim, and claims 7–12 depend therefrom. Claim 6 is reproduced below.

> 6. A keyfact-based text retrieving method comprising:
>
> keyfact extracting step for analyzing a document collection and a user query, and extracting keywords without part-of-speech ambiguity from said document collection and said user query, and respectively extracting keyfacts of said document collection and said user query from said keywords;
>
> keyfact indexing step for calculating the frequency of said keyfacts of said document collection and generating a keyfact list of said document collection for a keyfact index structure; and
>
> keyfact retrieving step for receiving said keyfact of said user query and said keyfacts of said document collection and defining a keyfact retrieval model in consideration of weigh factors according to a keyfact pattern and generating a retrieval result.

Ex. 1001, 9:54–10:7.

D. Instituted Grounds of Unpatentability

We instituted *inter* partes review based on the following grounds of unpatentability under 35 U.S.C. § 103(a)[2] as follows (Dec. 9–10, 38):

Claim(s) Challenged	35 U.S.C. §	Reference(s)/Basis
6–12	103(a)	Braden-Harder[3], Grossman[4]
6–12	103(a)	Braden-Harder, Heidorn[5], Messerly[6], Grossman
7–9	103(a)	Braden-Harder, Grossman, Kucera[7]
7–9	103(a)	Braden-Harder, Heidorn, Messerly, Grossman, Kucera
9	103(a)	Braden-Harder, Grossman, Miller[8]
9	103(a)	Braden-Harder, Heidorn, Messerly, Grossman, Miller
9	103(a)	Braden-Harder, Grossman, Kucera, Miller

[2] The Leahy-Smith America Invents Act, Pub. L. No. 112-29, 125 Stat. 284 (2011) ("AIA"), amended several provisions of 35 U.S.C., including § 103. Because the '908 patent has an effective filing date before the effective date of the applicable AIA amendments, we refer to the pre-AIA version of 35 U.S.C. § 103. Petitioner asserts, and Patent Owner does not dispute, that each relied upon reference is prior art under the pre-AIA version. Pet. 2, 7, 9, 12, 15–16; *see generally* PO Resp.

[3] US 5,933,822, issued Aug. 3, 1999 (Ex. 1020, "Braden-Harder").

[4] David A. Grossman and Ophir Frieder, *Information Retrieval: Algorithms and Heuristics*, Kluwer International Series in Engineering and Computer Science (Kluwer Academic Publishers, Springer Science + Business Media New York prtg.) (1st. ed. 1998) (Ex. 1010, "Grossman").

[5] US 5,966,686, issued Oct. 12, 1999 (Ex. 1022, "Heidorn").

[6] US 6,076,051, issued June 13, 2000 (Ex. 1025, "Messerly").

[7] US 4,868,750, issued Sept. 19, 1989 (Ex. 1011, "Kucera").

[8] George A. Miller et al., *Introduction to WordNet: An Online Lexical Database*, International Journal of Lexicography, Vol. 3, No. 4, 235–244 (1990) (Ex. 1027, "Miller").

Claim(s) Challenged	35 U.S.C. §	Reference(s)/Basis
9	103(a)	Braden-Harder, Heidorn, Messerly, Grossman, Kucera, Miller

II. DISCUSSION

A. Principles of Law

To prevail in its challenges to Patent Owner's claims, Petitioner must demonstrate by a preponderance of the evidence that the claims are unpatentable. 35 U.S.C. § 316(e); 37 C.F.R. § 42.1(d) (2019). A patent claim is unpatentable under 35 U.S.C. § 103(a) if the differences between the claimed subject matter and the prior art are such that the subject matter, as a whole, would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. *KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 406 (2007). The question of obviousness is resolved on the basis of underlying factual determinations including (1) the scope and content of the prior art; (2) any differences between the claimed subject matter and the prior art; (3) the level of ordinary skill in the art; and (4) when in evidence, objective evidence of nonobviousness.[9] *Graham v. John Deere Co.*, 383 U.S. 1, 17–18 (1966).

B. Level of Ordinary Skill

In determining the level of ordinary skill in the art, various factors may be considered, including the "type of problems encountered in the art; prior art solutions to those problems; rapidity with which innovations are made; sophistication of the technology; and educational level of active

[9] Patent Owner does not present any objective evidence of nonobviousness as to the challenged claims.

workers in the field." *In re GPAC Inc.*, 57 F.3d 1573, 1579 (Fed. Cir. 1995) (citation omitted). Petitioner relies on the declaration testimony of Dr. Bernard J. Jansen, who testifies that a person having ordinary skill in the art "would have had a Bachelor's degree in computer science, information technology, or the equivalent, and one to two years of experience working with at least one of information retrieval systems, the structure and organization of databases, and natural language processing." Pet. 22 (citing Ex. 1003 ¶ 142). According to Dr. Jansen, "significant work experience in any of these areas could substitute for formal education, and vice versa." Ex. 1003 ¶ 142.

Patent Owner indicates Petitioner's proposed assessment is improper because it does not define thresholds for "significant work experience" or "formal education." PO Resp. 4. Patent Owner, however, does not offer its own assessment of the level of skill in the art, or indicate that the resolution of any issue depends on a specific type of work experience or a specific level of education. *See id.* Patent Owner additionally states that even if Petitioner's proposed level of qualifications were applied, the Petition does not establish unpatentability. *Id.* at 5.

We adopt the assessment offered by Petitioner as it is consistent with the '908 patent and the asserted prior art. We further note that the prior art of record in the instant proceeding reflects the appropriate level of ordinary skill in the art. *Cf. Okajima v. Bourdeau*, 261 F.3d 1350, 1354–55 (Fed. Cir. 2001) (holding the Board may omit specific findings as to the level of ordinary skill in the art "where the prior art itself reflects an appropriate level and a need for testimony is not shown").

C. Claim Construction

In an *inter partes* review for a petition filed on or after November 13, 2018, "[claims] of a patent . . . shall be construed using the same claim construction standard that would be used to construe the [claims] in a civil action under 35 U.S.C. § 282(b), including construing the [claims] in accordance with the ordinary and customary meaning of such claims as understood by one of ordinary skill in the art and the prosecution history pertaining to the patent." *See* 37 C.F.R. § 42.100(b) (2019); *see also Phillips v. AWH Corp.*, 415 F.3d 1303, 1312–14 (Fed. Cir. 2005) (en banc).

"step for" limitations

Petitioner argues that the claim 6 "keyfact extracting step for," "keyfact indexing step for," and "keyfact retrieving step for" (collectively, "keyfact steps") should each be construed as "step-for" limitations in accordance with 35 U.S.C. § 112, ¶ 6. Pet. 25–31. For the same reasons presented in its Preliminary Response, Patent Owner argues that these claim 6 terms should not be construed in accordance with 35 U.S.C. § 112, ¶ 6. *Compare* PO Resp. 10–14, *with* Prelim. Resp. 15–19. Although we preliminarily agreed that the claims should not be construed in accordance with 35 U.S.C. § 112, ¶ 6, we also determined that whether the claim 6 terms invoke 35 U.S.C. § 112, ¶ 6 "is not material to resolving the issues currently before us, because Petitioner has shown a reasonable likelihood of prevailing under either construction." Dec. 17 (explaining how Petitioner's proposed construction is narrow, whereas Patent Owner's proposed construction is broad, and that Petitioner contends that the prior art renders the claims obvious "regardless of whether § 112, ¶ 6 applies"). Petitioner agrees with

our initial determination that we need not "address the issue because the claims are unpatentable even under Uniloc's argument." Pet. Reply 5.

Patent Owner, however, presents numerous arguments why the claims should not be construed in accordance with 35 U.S.C. § 112, ¶ 6 without clearly explaining why it is necessary to address § 112, ¶ 6 in order to resolve the controversies on the merits that are before us. PO Resp. 9–19; Sur-reply 7–9. As discussed below, beside the arguments directed to the sufficiency of the prior art teaching the "keyfact steps" limitations based on Patent Owner's construction of the term "keyfact," Patent Owner fails to explain sufficiently what is lacking from Petitioner's challenges and related analyses provided that assert that claims 6–12 would have been obvious based on the cited prior art *under either construction*. For example, Patent Owner argues that "Petitioner fails to recognize that the claim language recites various *acts*" and that Petitioner construes each recited act as a function. PO Resp. 9. In applying a 112, ¶ 6 analysis to claim 6, however, Petitioner looks to the '908 patent for the *corresponding acts* associated with the recited function. *See, e.g.*, Pet. 26–31. Thus, Patent Owner's arguments appear to us to be a matter of form over substance because Petitioner addresses the corresponding acts by looking to the Specification of the '908 patent. Tr. 8:12–25. Patent Owner does not persuade us otherwise. For example, Patent Owner argues that the claim 6 "keyfact extracting step" requires "a respective set of acts" such as "'analyzing …,' 'extracting …,' and 'respectively extracting … .'" PO Resp. 12–13 (alterations in original). Patent Owner argues that "Petitioner fails to comprehend the logical interrelationship between groups of *acts* and the underlying respective result, expressed as a 'step,' that each group of *acts* operated to collectively achieve." *Id.* at 14. Patent Owner, however, fails to identify or explain how

Petitioner's analysis of the claim language and prior art fails to address the claimed acts. As explained in detail below, we determine that Petitioner shows, by a preponderance of the evidence, how Braden-Harder and Grossman teach or suggest each of the claimed steps, such as the "keyfact extracting step." Pet. 33–47. As further explained below, Patent Owner does not present meaningful arguments to the contrary.

Patent Owner further argues that Petitioner's proposed constructions "violate[] both the doctrine of claim differentiation and the 'all elements' rule." PO Resp. 14–19. In particular, Patent Owner argues that dependent claim 7 recites additional acts required for the "keyfact extracting step" of claim 6 that are distinct from what claim 6 recites. *Id.* at 15–16. Patent Owner argues that "the 'analyzing' act of claim 6 (expressly directed to *both* 'a document collection *and* a user query' together) must require something more than the distinct 'analyzing' of dependent claim 7 (which is directed, instead, to 'morphology of an input sentence')." *Id.* at 15. Patent Owner contends that Petitioner erroneously conflates the acts of claim 6 with the acts of the dependent claims "thereby failing to give meaningful effect to those *acts* affirmatively recited in claim 6." *Id.* at 16–19.

Petitioner argues that the dependent claims "all identically recite '***wherein said step of … comprises the steps of …***' language." Pet. Reply 5 (alterations in original). As such, Petitioner argues that the dependent claims should be "construed as describing details required as part of performing the previously introduced steps" and that the dependent claims do "not add new 'distinct and additional' steps—contrary to Uniloc's contentions." *Id.* at 5–9 (citing *Realtime Data, LLC v. Iancu,* 912 F.3d 1368, 1375 (Fed. Cir. 2019)), 24–25 (arguing that dependent claims 7, 10, and 11 do not add new distinct and additional steps, but "give details of the more general steps previously

claimed"). As such, Petitioner argues that "Uniloc's arguments fall with its own erroneous interpretation of the dependent claims." *Id.* at 24–25. Petitioner further argues that whether 112, ¶ 6 applies "[t]he challenged claims are unpatentable either way." *Id.* at 8–9. Petitioner contends that the "all elements" rule is not applicable here, because that "rule relates to the doctrine of equivalents for infringement." *Id.* at 9 (citing *Warner-Jenkinson Co. v. Hilton Davis Chem. Co.*, 520 U.S. 17, 29 (1997)).

We agree with Petitioner that the dependent claims do not add new distinct steps or acts, but "give details of the more general steps previously claimed." *Id.* at 5–9, 24–25. That is, we agree with Petitioner that claim 7 recites "wherein said step of keyfact extracting comprises the steps of," which should be construed as describing details required as part of performing the previously introduced step. *Id.* at 5. On the other hand, Patent Owner fails to explain sufficiently why we should construe the dependent claims as reciting additional acts to the steps recited in claim 6. Accordingly, we agree with Petitioner that Patent Owner's arguments regarding claim differentiation necessarily fail because the arguments are based on an erroneous interpretation of the dependent claims to include additional steps or acts untethered from the steps recited in claim 6. Additionally, we agree with Petitioner that whether § 112, ¶ 6 applies "[t]he challenged claims are unpatentable either way." *Id.* at 8–9.[10]

[10] Patent Owner's arguments that Petitioner fails to address or rebut our preliminary determination that "claim 6 does not invoke 35 U.S.C. § 112, ¶ 6" (Sur-reply 7) are misplaced because, as discussed, whether the claim 6 terms invoke 35 U.S.C. § 112, ¶ 6 is not material to resolving the issues here.

Accordingly, consistent with the Decision to Institute, we determine that it is not necessary to construe the claim 6 "step for" terms because whether the claim 6 keyfact steps invoke 35 U.S.C. § 112, ¶ 6 is not material to resolving the issues currently before us because Petitioner has shown that the challenged claims are unpatentable under either construction. *See Vivid Techs., Inc. v. Am. Sci. & Eng'g, Inc.*, 200 F.3d 795, 803 (Fed. Cir. 1999) (holding that "only those terms need be construed that are in controversy, and only to the extent necessary to resolve the controversy"); *see also Nidec Motor Corp. v. Zhongshan Broad Ocean Motor Co. Matal*, 868 F.3d 1013, 1017 (Fed. Cir. 2017) (citing *Vivid Techs.* in the context of an *inter partes* review).

"keyfact"

Claim 6 recites the term "keyfact" several times. For example, claim 6 recites, "extracting keyfacts of said document collection and said user query from said keywords." Patent Owner argues that the term "keyfact" means "a factual extraction of a sentence which expresses semantic relation between words in the sentence in the form of [object, property]." PO Resp. 9 (citing Ex. 1001, 4:58–60). Patent Owner further argues that "[t]his form represents a paired relationship between an 'object' that is the head and a 'property' that is the modifier." *Id.* at 7 (citing Ex. 1001, 1:8–10). Petitioner argues that we need not construe the term "keyfact," because the challenged claims are unpatentable over the asserted prior art even under Patent Owner's construction. Pet. Reply 1–2.

For the reasons provided in this Decision discussed below, we agree with Petitioner that Braden-Harder's "extracted logical forms are each 'a factual extraction of a sentence which expresses semantic relation between words in the sentence in the form of [object, property].'" *See id.* at 10

(citing Pet. 47; Ex. 1003 ¶¶ 221–222). As explained below, we disagree, however, that a "keyfact" must be in the exact paired form "[object, property]" and nothing else. For purposes of this Decision, we need not otherwise expressly construe any claim terms.

D. Asserted Obviousness of Claims 6–12 over Braden-Harder, Heidorn, Messerly, and Grossman

1. Braden-Harder

Braden-Harder describes an information retrieval system that utilizes natural language processing to process results retrieved by an information retrieval search engine that ultimately yields a set of retrieved documents. Ex. 1020, code (57). Each document is subjected to natural language processing to produce a set of logical forms, where each logical form encodes, in a word-relation-word manner, semantic relationships between words in a phrase. *Id.* A user-supplied query is analyzed in the same manner to yield a set of corresponding logical forms, where documents are ranked as a predefined function of the logical forms from the documents and the query. *Id.*

A high-level block diagram of an information retrieval system is illustrated in Figure 1, reproduced below.

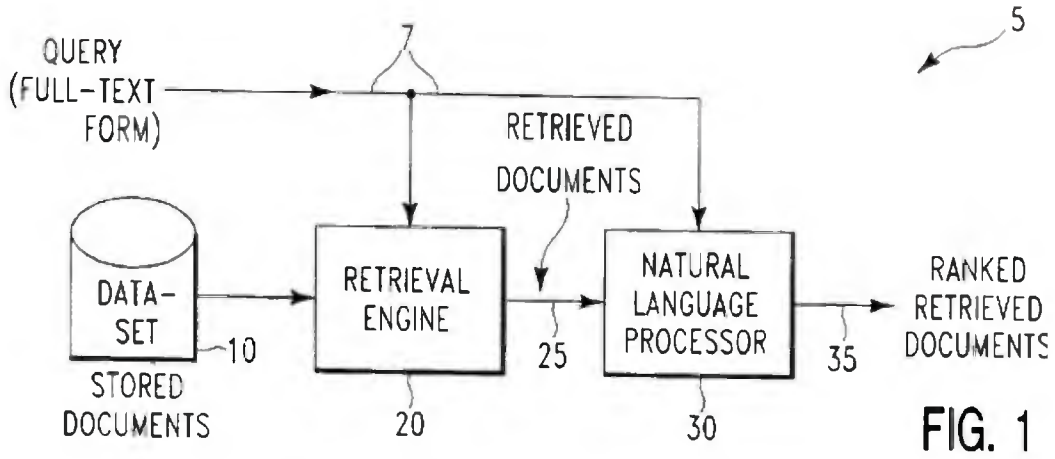


FIG. 1

Figure 1 is a high-level block diagram of an information retrieval system 5. *Id.* at 7:25–27. In operation, a user supplies a search query to system 5. *Id.* at 7:35–36. System 5 applies this query both to information retrieval engine 20 and natural language processor 30. *Id.* at 7:40–41. In response to the query, engine 20 searches through a dataset 10 of stored documents to yield a set of retrieved documents. *Id.* at 7:41–43. The set of documents is then applied, as symbolized by line 25, as an input to processor 30. *Id.* at 7:43–46. Within processor 30, each of the documents in the set is subjected to natural language processing, specifically morphological, syntactic and logical form, to produce logical forms for each sentence in that document. *Id.* at 7:46–50. Each such logical form encodes semantic relationships between words in a linguistic phrase in that sentence. *Id.* at 7:50–53. Processor 30 analyzes the query in an identical fashion to yield a set of corresponding logical forms. *Id.* at 7:53–55. Processor 30 then compares the set of forms for the query against the sets of logical forms associated with each of the documents in the set in order to ascertain any match between logical forms in the query set and logical forms for each document. *Id.* at 7:55–59. Documents that produce no matches are

eliminated from further consideration, where each remaining document that contains at least one match is retained and heuristically scored by processor 30. *Id.* at 7:59–63. Finally, processor 30 presents the retained documents to the user rank-ordered based on their respective scores. *Id.* at 8:2–4.

2. Heidorn

Heidorn describes a method and system for performing semantic analysis of an input sentence within a natural language processing ("NLP") system. Ex. 1022, 6:28–30. A semantic analysis subsystem receives a syntax parse tree generated by morphological and syntactic subsystems. *Id.* at 6:30–32. The semantic analysis subsystem applies two sets of semantic rules to make adjustments to the received syntax parse tree. *Id.* at 6:32–34. The semantic analysis subsystem then applies a third set of semantic rules to create a skeletal logical form graph from the syntax parse tree. *Id.* at 6:34–36. The semantic analysis subsystem finally applies two additional sets of semantic rules to the skeletal logical form graph to provide semantically meaningful labels for the links of the logical form graph, to create additional logical form graph nodes for missing nodes, and to unify redundant logical form graph nodes. *Id.* at 6:36–42. The final logical form graph generated by the semantic analysis subsystem represents the complete semantic analysis of an input sentence. *Id.* at 6:42–44.

An example semantic analysis subsystem is illustrated in Figure 43, reproduced below.

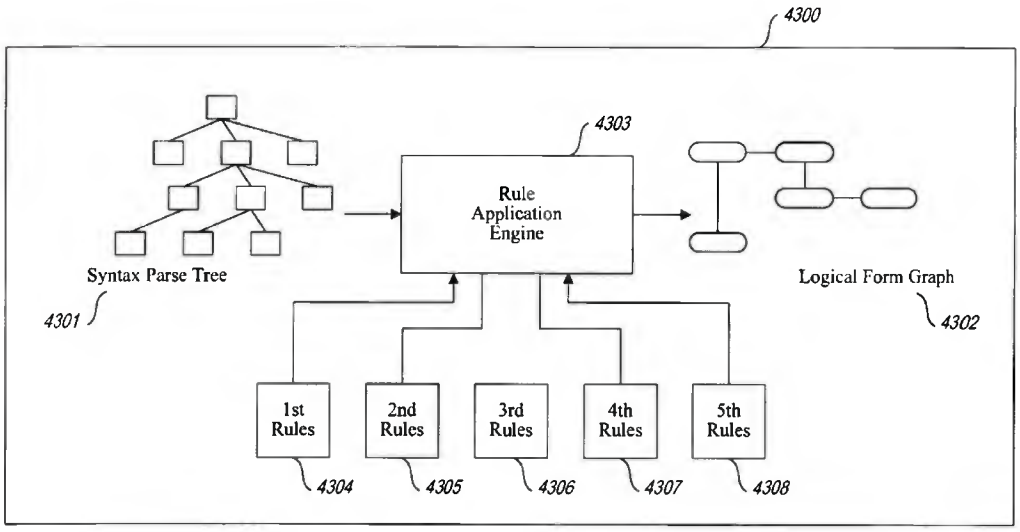


Fig. 43

Figure 43 is a block diagram of a computer system for a semantic analysis subsystem. *Id.* at 13:49–50. The computer 4300 contains a memory with semantic rules 4304–4308 and rule application engine 4303. *Id.* at 13:50–51. The rule application engine, under control of a central processing unit, applies the five sets of rules to a syntax parse tree 4301 to generate a corresponding logical form graph 4302. *Id.* at 13:51–54. The syntax parse tree is generated by the morphological and syntactic subsystems, which are not shown. *Id.* at 13:55–56.

3. Messerly

Messerly describes performing information retrieval using semantic representation of text. Ex. 1025, code (57). A tokenizer generates, from an input string information retrieval, tokens that characterize the semantic relationship expressed in the input string. *Id.* The tokenizer first creates,

from the input string, a primary logical form characterizing a semantic relationship between the selected words in the input string. *Id.* The tokenizer then identifies hypernyms that each have an "is a" relationship with one of the selected works in the input string. *Id.* The tokenizer then constructs one or more alternative logical forms from the primary logical form. *Id.* For each of one or more of the selected words in the input string, the tokenizer constructs each alternative logical form by replacing the selected word in the primary logical form with an identified hypernym of the selected word. *Id.* Finally, the tokenizer generates tokens representing both the primary logical form and the alternative logical forms. *Id.*

An example overview flow diagram showing steps performed in order to construct and access an index semantically representing target documents is illustrated in Figure 3, reproduced below.

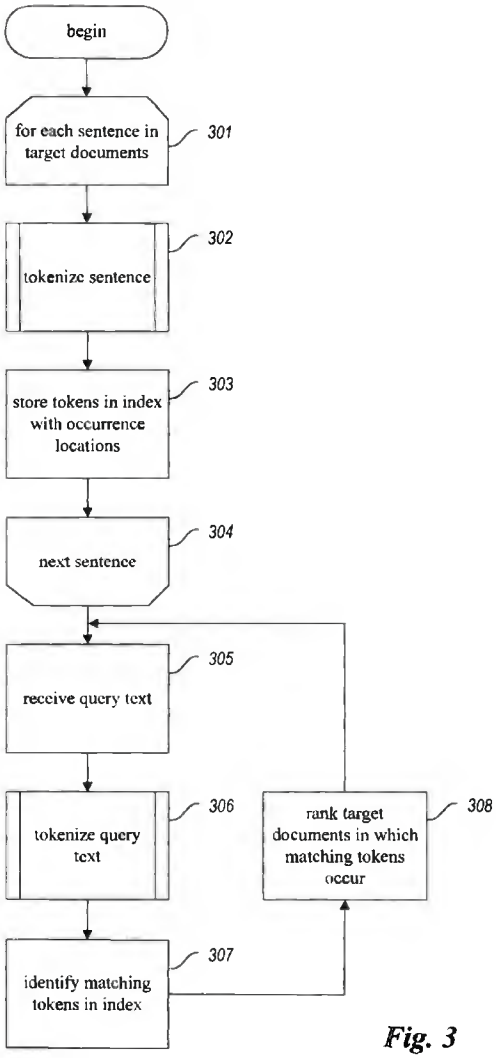


Fig. 3

Figure 3 is an overview flow diagram showing the steps performed by a facility in order to construct and access an index semantically representing target documents. *Id.* at 5:31–34. As represented in steps 301–304, the facility: converts each sentence of each target document into a number of tokens representing an expanded logical form portraying the relationship between the important words in the sentence (including hypernyms having similar meanings); and stores the semantic tokens in an index along with a location in the target documents where the sentence occurs. *Id.* at 5:34–41,

5:56–59. After all of the target documents have been indexed, the facility is able to process information retrieval queries against the index. *Id.* at 5:41–44. As represented in steps 305–308, for each received query, the facility: tokenizes the text of the query (i.e., converts the sentence into semantic tokens together representing an expanded logical form for the query text); compares these semantic tokens to the semantic tokens stored in the index to identify locations in the target documents for which these semantic tokens have been stored; and ranks the target documents containing these semantic tokens in the order of their relevance to the query. *Id.* at 5:44–53.

4. *Grossman*[11]

Grossman describes the retrieval of information (e.g., finding relevant documents) in response to user queries. Ex. 1010, 1. More specifically, Grossman describes building an inverted index as part of an information retrieval system to find quickly terms in a document collection. *Id.* at 134. As described in Grossman, the inverted index includes a set of files, including: an index file that contains an actual posting list for each distinct term in the collection; a document file that contains information about each distinct document; and a weight file that contains the weight for each document. *Id.* at 136–137.

Grossman describes a preprocessor that reads an input file, outputs separate files, and bulk-loads the output files into a relational database. *Id.*

[11] Petitioner contends that Grossman (Exhibit 1010) is a textbook published in 1998, and that a copy of the textbook was cataloged and available at the UCB Library by November 17, 1998. Pet. 12 (citing Ex. 1009). Patent Owner does not challenge Petitioner's contentions as to the prior art status of Grossman. *See generally* PO Resp. For the reasons provided in the Petition, we agree with Petitioner that Grossman is "prior art at least under §102(a)." *Id.*

at 169–170. Each output file corresponds to a relation. *Id.* at 170. The first relation, "DOC," contains information about each document. *Id.* The second relation, "INDEX," models the inverted index and indicates which term appears in which document and how often the term has appeared. *Id.* A third relation, "TERM," tracks statistics for each term based on its number of occurrences across the document collection. *Id.*

Example DOC, INDEX, and TERM relations are illustrated in Tables 5.5, 5.6, and 5.8, reproduced below.

Table 5.5. DOC

DocId	DocName	PubDate	Dateline
1	WSJ870323-0180	3/23/87	TURIN, Italy
2	WSJ870323-0161	3/23/87	Du Pont Company, Wilmington, DE

Table 5.6. INDEX

DocId	Term	TermFrequency
1	commercial	1
1	vehicle	1
1	sales	1
1	italy	1
1	february	1
1	year	1
1	according	1
...	...	...
2	krol	2
2	president	2
2	diversified	1
2	company	1
2	succeeding	1
2	dale	1
2	products	2
...	...	...

Table 5.8. TERM

Term	Idf
according	0.9031
commercial	1.3802
company	0.6021
dale	2.3856
diversified	2.5798
february	1.4472
italy	1.9231
krol	4.2768
president	0.6990
products	0.9542
...	...
...	...
sales	1.0000
succeeding	2.6107
vehicle	1.8808
year	0.4771
...	...

Table 5.5 illustrates a DOC relation which models the documents in the document collection; Table 5.6 illustrates an INDEX relation which models an inverted index by storing the occurrences of a term in a document; and Table 5.8 illustrates a TERM relation which contains an inverse document frequency ("idf") of a term, where idf = log (number of documents / number of documents which contain the term). *Id.* at 16, 172–175.

5. Discussion

Petitioner contends that claims 6–12 are unpatentable under 35 U.S.C. § 103(a) as obvious over Braden-Harder and Grossman, or Braden-Harder, Heidorn, Messerly, and Grossman. Pet. 33–84. For the Braden-Harder and Grossman challenge, Petitioner contends that Braden-Harder "refers to and incorporates two co-pending U.S. patent applications assigned to Microsoft," which are the Heidorn and Messerly patent applications. *Id.* at 5–6 (citing Ex. 1020, 14:53–61). Petitioner further explains that, because Braden-

Harder identifies the Heidorn and Messerly patent applications and where they can be found, "such broad and unequivocal language is plainly sufficient to incorporate both patent applications in their entirety." *Id.* at 6 (citing *Paice LLC v. Ford Motor Co.*, 881 F.3d 894, 907 (Fed. Cir. 2018)). Petitioner further contends that "[e]verything disclosed in the Heidorn and Messerly patent applications is 'effectively part of' Braden-Harder as if these disclosures were 'explicitly contained therein.'" *Id.* Petitioner explains that the Petition cites to the Messerly and Heidorn patents instead of the respective patent applications because the respective patents "disclose substantively the same subject matter and contain substantively identical disclosures." *Id.* at 7 n.2, 9 n.3.

For the Braden-Harder, Heidorn, Messerly, and Grossman challenge, Petitioner argues that to the extent we decide that the disclosures in Heidorn and Messerly were not incorporated into Braden-Harder, it would have been obvious to combine these references. Pet. 33 (citing *Commonwealth Sci. & Indus. Research Org. v. Buffalo Tech. (USA), Inc.*, 542 F.3d 1363, 1372 (Fed. Cir. 2008)).

Patent Owner does not dispute Petitioner's assertions that the Heidorn and Messerly applications are incorporated by reference into Braden-Harder, or that the respective Heidorn and Messerly patents "disclose substantively the same subject matter and contain substantively identical disclosures" as the incorporated Heidorn and Messerly applications. *See generally* PO Resp. Nor does Patent Owner contest Petitioner's alternative theory, that it would have been obvious to combine Braden-Harder, Heidorn, Messerly, and Grossman. *Id.* For purposes of this Decision, we agree with Petitioner, for the reasons provided in the Petition, that Heidorn and Messerly were properly incorporated into Braden-Harder at the time of the Braden-Harder

invention and that the respective Heidorn and Messerly patents disclose substantively the same subject matter and contain substantively identical disclosures.

Accordingly, we focus on Petitioner's challenge that claims 6–12 are unpatentable under 35 U.S.C. § 103(a) as obvious over Braden-Harder (incorporating Heidorn and Messerly) in view of Grossman.[12] In support of its showing, Petitioner relies upon the declaration of Dr. Bernard J. Jansen. *Id.* (citing Ex. 1003). Patent Owner contends that the combined teachings of Braden-Harder and Grossman do not teach the following limitations: (1) "keyfact"; (2) the acts for the "keyfact extracting step"; and (3) the "keyfact retrieving step." PO Resp. 10–28; Sur-reply 10–22. Patent Owner does not introduce testimony from a declarant in support of its positions.

For the reasons that follow, we conclude that Petitioner has met its burden of proving by a preponderance of the evidence that each of the challenged claims would have been obvious in view of the asserted prior art.

a. Claim 6: "a keyfact-based text retrieving method comprising:" (preamble)[13]

For the above limitation, Petitioner contends, and we agree, that Braden-Harder's information retrieval (IR) method describes subjecting documents and queries to natural language processing (NLP) to produce logical form triples, with each of the triples being of the form "word-relation-word." Pet. 33–34 (citing Ex. 1020, 5:22–24, 5:29–35, 11:44–55). Petitioner further contends, and we agree, that the logical-form triples

[12] We refer to this challenge as Braden-Harder in view of Grossman or Braden-Harder (incorporating Heidorn and Messerly) in view of Grossman.
[13] We need not determine whether the preamble is limiting because Petitioner shows that Braden-Harder meets the preamble.

express semantic relationships between important words in an input string or phrase. *Id.* at 34 (citing Ex. 1020, 5:18–20, 11:44–46; Ex. 1001, 1:53–59). Petitioner further contends that Braden-Harder's logical-form triples are equivalent to the claimed keyfacts and provide a concept-based retrieval. *Id.* (citing Ex. 1003 ¶ 176). Although Patent Owner argues that Petitioner fails to show that Braden-Harder in combination with Grossman meets the "keyfact extracting step," as explained below, Patent Owner does not dispute Petitioner's showing with respect to the preamble. *See generally* PO Resp.

b. Claim 6: "keyfact extracting step for analyzing a document collection and a user query, and extracting keywords without part-of-speech ambiguity from said document collection and said user query, and respectively extracting keyfacts of said document collection and said user query from said keywords"

For the "analyzing a document collection and a user query" portion of the above limitation, Petitioner contends that Braden-Harder subjects documents and user queries to morphological analysis to parse text into constituent words using a stored lexicon. Pet. 37–38 (citing Ex. 1020, 5:33–35, 11:62–12:1, 12:12–16, 12:30–34; Ex. 1003 ¶¶ 184–187; Ex. 1022, 1:22–24; Ex. 1029, 112). Petitioner further contends, and we agree, that a POSITA[14] would have understood that Braden-Harder's lexicon is a machine-readable dictionary. *Id.* at 38 (citing Ex. 1003 ¶¶ 184–187; Ex. 1029, 112). As pointed out by Petitioner, Braden-Harder's lexicon "includes various classes of words, such as e.g., prepositions, conjunctions, verbs, nouns, operators and quantifiers that define syntactic and semantic properties inherent in the words in an input string." Ex. 1020, 12:12–16. Petitioner further contends that using its lexicon, Braden-Harder's

[14] A person of ordinary skill in the art.

morphological analysis normalizes differing word forms to a common morphological form for use by a parser. *Id.* at 38 (citing Ex. 1020, 12:34–37).

Petitioner contends that to accomplish the function of "extracting keywords without part-of-speech ambiguity from said document collection and said user query," Braden-Harder teaches applying parts-of-speech tags to each extracted word. *Id.* at 39. Specifically, Petitioner asserts that Braden-Harder's morphological analysis attributes each word, or phrase (noun phrase) in an input string, with a category of its corresponding grammatical function (part-of-speech tag for that word). *Id.* (citing Ex. 1020, 12:40–45). Petitioner contends that a noun extracted from a noun phrase in Braden-Harder is a keyword used to retrieve documents. *Id.* (citing Ex. 1003 ¶ 189; Ex. 1022, 3:36–40).

Petitioner further contends that Braden-Harder's system syntactically analyzes stemmed words using grammatical rules and attributes to yield a syntactic parse tree with at least three levels (Table 1): top (root) node (e.g., declarative sentence), intermediate level nodes (e.g., verb, noun, or prepositional phrases), and a lower-level (leaf) nodes which represent the initial set of attribute/value records. *Id.* at 39–40 (citing Ex. 1020, 12:37–40, Table 1; Ex. 1022, 3:54–59). Petitioner asserts that Table 1 of Braden-Harder shows that the intermediate-level nodes and lower-level nodes represent part-of-speech tag sequences for the input string. *Id.* at 40 (citing Ex. 1020, Table 3). Petitioner asserts, and we agree, that in Braden-Harder's syntactic-parse tree, the extracted words and phrases have no "part-of-speech ambiguity." *Id.*

Petitioner contends that Heidorn describes the specific details of syntactic processing (extracting keywords without part-of-speech

ambiguity). *Id.* at 40–42 (citing Ex. 1022, 4:2–3, 4:44–51, Figs. 6, 22; Ex. 1003 ¶ 199). Petitioner argues that, to the extent the claim requires replacing a sequence of tags rather than selecting a final sequence of tags, like the '908 patent's "KEY," Braden-Harder similarly generates an intermediate-level tag sequence for an input sentence in which a sequence of lower-level-tagged noun phrases are replaced by a single tag "NP," resulting in a final sequence tag at the intermediate-level sequence (i.e., NP-VERB-NP). *Id.* at 42–43 (citing Ex. 1020, 12:40–45, 12:66–13:10, Table 1).

As further part of the "keyfact extracting step for" limitation, claim 6 recites, "respectively extracting keyfacts of said document collection and said user query from said keywords." Petitioner contends that Braden-Harder teaches processing the syntactic parse tree using a set of semantic rules ("keyfact pattern rules") to yield a logical form graph ("keyfact pattern") and incorporates Heidorn for the specific details of the logical form processing. *Id.* at 43 (citing Ex. 1020, 14:50–51, Figs. 5A–5D). Petitioner asserts that, like the '908 patent's disclosure, the final tag sequence is applied to a keyfact pattern rule to extract a keyfact pattern (Ex. 1001, 6:33–36), and Heidorn similarly applies semantic rules to the tag sequence to generate a logical form graph. *Id.* at 44–45 (citing Ex. 1022, 2:22–30, Fig. 43; Ex. 1003 ¶ 209). Petitioner further asserts that the extracted logical form graph captures the structure of arguments and adjuncts for each input string and provides a pattern representing the sentence in logical form. *Id.* at 45 (citing Ex. 1022, 13:56–58; Ex. 1003 ¶ 210). Petitioner contends, and we agree, that Braden-Harder applies rules to the logical form graph to generate a list of logical form triples, which are extracted based on semantic relationships of the nouns (keywords) with other words in the sentence. *Id.*

at 45–47 (citing Ex. 1020, 14:3–4, 14:18–24, 14:28–33, 14:37–41, Figs. 5B–5D; Ex. 1003 ¶¶ 211–219).

Important to this Decision, Petitioner argues, and we agree, that to the extent that "extracting keyfacts of said document collection and said user query from said keywords" requires generating keyfacts that have forms of [object, property], "a POSITA would have understood that Braden-Harder generates logical form triples for 'the case of a constituent with noun modifiers' (*see* Table 3), such as 'Nadj' and 'Mods' relations, which exhibit the semantic relationship of the form [object, modifier]." *Id.* at 47.[15] Alternatively, Petitioner contends, and we agree, that it would have been obvious to generate other logic forms that do not include relations to enable structural paraphrasing. *Id.* (citing Ex. 1020, 11:51–55, 25:49–60). Petitioner further contends, and we agree, that a POSITA would have been motivated "to generate logic forms that do not include the relations for 'noun appositive' or 'relative clause' constructions to increase matches with 'sufficiently similar semantic content.'" *Id.* (citing Ex. 1020, 25:51–52; Ex. 1003 ¶¶ 221–222).

[15] During the oral hearing, Petitioner stated, "[t]he difference between property and modifier is inconsequential. We can see from the '908 patent at [column 1, lines 6–12] . . . that the property is the modifier, those are the same things so that is a distinction without a difference." Tr. 17:14–18. We agree that the property described in the '908 patent is the same as a modifier because the '908 patent describes it as such. Ex. 1001, 1:7–12 ("In particular, the methods describe the formalized concept of a document as a pair comprising an object that is the head *and a property that is the modifier*, and uses the information described by the pair as index information for efficient document retrieval" (emphasis added)). Patent Owner does not contend otherwise. *See generally* PO Resp.

Patent Owner argues that the "Petition fails to prove obviousness of claim limitations reciting the 'keyfact' term, under a proper interpretation that accurately reflects the intrinsic evidence." PO Resp. 20. Patent Owner further argues that Petitioner focuses on Braden-Harder in addressing "analyzing a document collection and a user query" by alleging that Braden-Harder discloses using a lexicon to recover various classes of words. *Id.* at 20–21. Patent Owner argues that "a mere distinction of 'classes or words' . . . is distinguishable on its face from the distinct concept of a 'keyfact' that is an extracted expression of fact, much less one that represents a semantic relation between words of a sentence in the form of [object, property], as disclosed and claimed in the '908 patent." *Id.* at 21. Patent Owner argues that in the '908 patent, the class of noun determines the object or the property of the keyfacts and that the "Braden-Harder scheme makes no determination between a class of noun, much less one used to generate a 'keyfact' having the form [object, property]." *Id.* at 22. Patent Owner further contends that Braden-Harder's triples are distinguishable from the keyfact extraction disclosed and claimed, because "the '908 patent disparages the sort of phrase-based scheme Braden-Harder employs as being subject to certain technical disadvantages." *Id.* (citing Ex. 1001, 1:19–28, 1:42–45).

Petitioner responds that Braden-Harder meets the claim limitation under Patent Owner's construction for "keyfact" presented in its Response. Pet. Reply 10 (citing Pet. 47; Ex. 1003 ¶¶ 221–222). Petitioner argues, and we agree, that the Petition explains that Braden-Harder teaches various special graph walks, including one for the case of a constituent with noun modifiers. *Id.* (citing Pet. 45; Ex. 1020, 14:37–41, Fig. 5D; Ex. 1003

¶¶ 216–217). Petitioner points out that the Petition specifically explains that if keyfacts were construed to have the form [object, property], a person having ordinary skill in the art would have understood that Braden-Harder generates logical form triples for the case of a constituent with noun modifiers such as "Nadj" and Mods" relations which exhibit the semantic relationship of the form [object, modifier]. *Id.* at 11 (citing Pet. 47; Ex. 1003 ¶ 221). In particular, Dr. Jansen testifies that:

> [A] person of ordinary skill in the art would have understood that Braden-Harder generates logical form triple for "the case of a constituent with noun modifiers" (*see* Table 3), such as "Nadj" and "Mods" relations that exhibit the semantic relationship of the form [object, Mods, modifier]. Because Braden-Harder generates logical form triples for noun modifiers, Braden-Harder generates keyfacts having the forms of [object, modifier].

Ex. 1003 ¶ 221. Dr. Jansen further testifies, "each of Braden-Harder's logical form triples 'contains two node words … linked by a semantic relationship there between." *Id.* ¶ 176 (alteration in original).[16] We give substantial weight to Dr. Jansen's unrebutted testimony because it is supported by the disclosures in Braden-Harder. In particular, Braden-Harder's triples are in the form [object, property] because they are in the form, for example, [object, Mods, modifier[17]], where each triple contains

[16] The Petition lists eight examples of Braden-Harder's triples that meet the form [object, property]. Pet. 45–46; Ex. 1003 ¶¶ 216–221; *see also* Tr. 18:18–24 (explaining that in the example "BOWL – Mods – SHARK" from Figure 5D, "that 'BOWL' would be the object and 'SHARK' would be the modifier for the property").

[17] As explained above in footnote 15, the record supports that there is no distinction between modifier and property and we treat them the same, as do the parties.

two node words linked by a semantic relationship. *See* Ex. 1001, 1:7–12. We agree with Petitioner that "[i]n the case of noun modifier, Braden-Harder's logical form triples represent 'a paired relationship' between an object that is the noun and a property that is the modifier, and are thus in the paired form of [object, property]." Pet. Reply 12, 18 (citing Ex. 1001, 1:7–12).

Patent Owner argues that the use of triples in Braden-Harder is distinguishable from claim language where keyfacts are extracted in a paired form, i.e. [object, property]. PO Resp. 22. We agree, however, with Petitioner's assertion that there is nothing in the '908 patent "specification or the prosecution history that would limit 'keyfacts' to at most two elements (e.g., object and property) and nothing more." *Id.* at 18–19. Patent Owner directs us to no intrinsic or extrinsic evidence to support its contention that a "keyfact" must be in the exact paired form "[object, property]" and nothing else. Indeed, as pointed out by Petitioner, Table 1 of the '908 patent includes, under the "keyfact term list" column "[KEY2 KEY1, NIL]." Ex. 1001, 6:18, Table 1.

In the Sur-reply, Patent Owner argues that "every single 'keyfact' example disclosed in the specification expresses semantic relation between words in the paired form of [object, property]." Sur-reply 2 (citing Ex. 1001, code (57), Table 1, 1:8–10, 1:16–18, 4:58–60, 6:15–30, 6:38–44) (emphasis omitted). Patent Owner argues that, in the example from Table 1 of [KEY2 KEY1, NIL], "the 'property' portion is 'NIL' or empty, such that KEY2 and KEY1 pair are the only two nonempty keyfact tags." *Id.* at 13 (emphasis omitted). We are not persuaded by Patent Owner's arguments that the '908 patent includes a "disavowal expressed in the statement 'All keyfacts express semantic relation between words in the form of [object,

property],'" such that a keyfact must contain a pair and nothing else. *Id.* at 12–13. We agree with Petitioner that the description from the '908 patent (all keyfacts express semantic relation between words in the form of [object, property]) does not preclude, for example, "more than the pair of node words," such as Braden-Harder's triples that include two node words in the form of [object, property] linked by a semantic relationship. Ex. 1003 ¶¶ 176, 221. Patent Owner responds that in every instance shown in Table 1, "the example keyfacts listed never consist of more than a pair of nonempty tags." Sur-reply 13 (emphasis omitted). However, even the form [KEY2 KEY1, NIL], for example, contains an indicator "NIL" along with two other terms (KEY2 and KEY1), resulting in a representation that is more than just a pair of terms, like Braden-Harder. Moreover, Patent Owner's argument that Braden-Harder's triples are "the sort of phrase-based scheme" that the '908 patent disparages is unavailing. PO Resp. 22 (citing Ex. 1001, 1:19–28, 1:42–45). As explained above, Braden-Harder applies rules to extract logical form triples, where each triple "is a factual extraction that represents the semantic relation between words of a sentence," which is distinguishable from phrase-based methods that use "purely syntactic" approaches. Pet. Reply 21–22 (citing Pet. 1, 34; Ex. 1003 ¶¶ 175–176; Ex. 1020, 3:46–59, 4:18–20, 4:61–67) (emphasis omitted).

We also are not persuaded by Patent Owner's arguments that claim 6 requires a "determination between a class of noun" used to generate a keyfact having the form [object, property]. PO Resp. 21–22 (citing Ex. 1001, 5:35–44). Claim 6 does not require making a determination between a class of noun. A class of noun is not in the claim language at all. Moreover, we agree with Petitioner that Braden-Harder nonetheless makes such a determination between a class of noun, and converts a sequence of nouns

into a single tag, like the '908 patent. Pet. Reply 16–18 (citing Pet. 42, 45–46; Ex. 1022, Fig. 22; Ex. 1020, 14:37–41; Ex. 1003 ¶¶ 199–200, 216–217).

Lastly, Patent Owner argues that for claim 6, Petitioner misconstrues acts as underlying functions and that "[i]t is significant here that claim 6 recites the 'keyfact extracting step' in the context of both a *document collection* (i.e., more than one document) *and* a 'user query' together, for each one of the corresponding acts." PO Resp. 23–24. Patent Owner also argues that Petitioner makes no distinction between its challenge of claim 6 "and the distinct and additional requirement" recited in dependent claim 7. *Id.* (emphasis omitted). For the reasons explained above in the claim construction section of this Decision, we disagree with Patent Owner that the dependent claims, such as claim 7, include additional acts, and instead we agree with Petitioner that the dependent "claims should be properly construed as describing details required as part of performing the previously introduced steps." Pet. Reply 5. Also, as explained above, Petitioner addresses the acts that Patent Owner argues are claimed in claim 6 by construing the claim under 35 U.S.C. § 112, ¶ 6.

To the extent Patent Owner contends that the Petition fails to show how Braden-Harder teaches or suggests the "keyfact extracting step" for *both a document collection* and a user query together, Patent Owner's argument is without merit. PO Resp. 23–24. First, the claim "keyfact extracting step" does not recite that the analyzing of a document collection and a user query be performed "together" or at the same time, as Patent Owner appears to imply. There simply is no such temporal requirement and Patent Owner fails to explain why we should construe the claim to include one. Moreover, Patent Owner fails to identify what is missing from Petitioner's showing. *Id.* at 23–24. For example, the Petition shows, and we

agree, that "the 'keyfact extracting step' is performed in the context of both a document collection and a user query." Pet. Reply 25–27 (citing Pet. 35–37; Ex. 1003 ¶¶ 180–183; Ex. 1020, Figure 10A); Pet. 33–34 (citing Ex. 1020, 5:22–24, 5:29–35, 11:44–55, and arguing "Braden-Harder's IR method subjects both documents and queries to NLP to produce 'a list of logical form triples, with each of the triples being illustratively of a form 'word-relation-word'"). Indeed, Braden-Harder describes that *each document* is "subjected to natural language processing, specifically morphological, syntactic and logical form, to ultimately produce appropriate logical forms for each sentence in each document. A user-supplied query is analyzed in the same manner to yield a set of corresponding logical form triples therefor." Ex. 1020, 5:29–35; *see also id.* at 26:33–61 (claim 1 reciting the same). Petitioner explains, and we agree, that Braden-Harder teaches a document indexing engine that performs a triple-generation process to extract logical-form triples *for each document being indexed*. Pet. 35 (citing Ex. 1020, 6:55–57, 20:62–21:4, Figs. 10A, 10B, 11; Ex. 1003 ¶ 180). Petitioner further explains that Braden-Harder performs a similar retrieval process *for analyzing a user query to yield logical form triples*. *Id.* at 36 (citing Ex. 1020, 15:18–21, Fig. 12A). Patent Owner directs us to no supporting evidence that rebuts Petitioner's showing, which we determine is sufficient. For all of the above reasons, we determine that Petitioner has shown sufficiently that Braden-Harder in view of Grossman teaches or suggests the disputed "keyfact" limitation.

As explained above, Petitioner alternatively contends, and we agree, that as to the claimed "keyfact" limitation, it would have been obvious to generate other logic forms that would not include relations to enable structural paraphrasing. Pet. 47 (citing Ex. 1020, 11:51–55, 25:49–60).

Petitioner further contends, and we agree, that a POSITA would have been motivated "to generate logic forms that do not include the relations for 'noun appositive' or 'relative clause' constructions to increase matches with 'sufficiently similar semantic content.'" *Id.* (citing Ex. Ex. 1020, 25:51–52; Ex. 1003 ¶¶ 221–222).[18] Braden-Harder describes that its "invention can readily utilize any other form, such as those noted above, that can portray a semantic relationship between words." Ex. 1020, 11:51–55. Braden-Harder further describes that "the criteria for determining a match, for purposes of identifying sufficiently similar semantic content across triples, *can be relaxed* to encompass paraphrases as matching." *Id.* at 25:50–53 (emphasis added). "A structural paraphrase is exemplified by use of either a noun appositive or a relative clause." *Id.* at 25:53–57. Dr. Jansen testifies that Braden-Harder's descriptions would have led a POSITA to "be motivated to generate logic forms that do not include the relations for 'noun appositive' or 'relative clause' constructions to increase matches with 'sufficiently similar semantic content'" thereby ensuring "that all relevant documents are retrieved from the document collection." Ex. 1003 ¶ 221. We give substantial weight to Dr. Jansen's testimony and agree with Petitioner's contentions that it would have been obvious to generate other logic forms that do not include relations to enable structural paraphrasing. Pet. 47.

In the Patent Owner Response, Patent Owner did not address Petitioner's alternative theory of how the prior art renders obvious the disputed "keyfact" limitation, but only made arguments regarding the theory in the Sur-reply. *See generally* PO Resp.; Pet. Reply 19–20; Sur-reply 14–

[18] During oral hearing, Petitioner stated that by not including the relation "it gets rid of that mod that's in the middle of the tuple" resulting in "literally . . . this exact form of object and property." Tr. 19:6–17.

16 ("Google also falsely suggests it is undisputed that it would have been obvious to modify Braden-Harder to *exclude* from its 'triples' the interposing relation label"). In other words, Patent Owner waited until the Sur-reply to address Petitioner's alternative showing made in the Petition. We determine it is too late for Patent Owner to do so. In the Scheduling Order, we advised Patent Owner "that any arguments not raised in the [Patent Owner Response] may deemed waived." Paper 15, 8; *see also In re NuVasive, Inc.*, 842 F.3d 1376, 1381 (Fed. Cir. 2016) (holding that patent owner's failure to proffer argument at trial as instructed in the scheduling order constitutes waiver). A sur-reply that belatedly raises a new issue may not be considered. *See, e.g.*, *Werner Co. v. Louisville Ladder, Inc.*, IPR2019-00336, Paper 34 at 39 (PTAB Aug. 24, 2020) (Final Written Decision) ("Patent Owner's arguments made for the first time in its Sur-reply are waived because they were not made in the Patent Owner Response, and Petitioner relied on their absence to its detriment and was deprived of the chance to present arguments and evidence in its Reply to rebut Patent Owner's assertions."); *see also* Consolidated Trial Practice Guide, 74 (November, 2019)[19]. Additionally, Patent Owner provides no explanation for presenting its arguments late. *See* Sur-reply 14–16.[20] Thus, for the reasons provided in the Petition, we determine that Petitioner's unrebutted

[19] Available at https://www.uspto.gov/sites/default/files/documents/tpgnov.pdf?MURL=.

[20] Even if we considered Patent Owner's untimely argument that Petitioner's proposed modification would "render Braden-Harder inoperable for its intended purpose," the underlying contentions are unsupported by evidence. Sur-reply 14–16. As stated above, Patent Owner does not introduce testimony from a declarant in support of its positions. Nor did Patent Owner cross-examine Dr. Jansen. Thus, Dr. Jansen's testimony stands unrebutted.

alternative showing is sufficient to show that the disputed limitation with the "keyfact" term would have been obvious.

c. Claim 6: "keyfact indexing step for calculating the frequency of said keyfacts of said document collection and generating a keyfact list of said document collection for a keyfact index structure"

For the above limitation, Petitioner contends that Braden-Harder (incorporating Messerly) in combination with Grossman teach keyfact indexing by calculating keyfact frequencies (tf) and document frequencies (df) to generate tables used to construct a keyfact index structure. Pet. 47–48. In particular, Petitioner contends, and we agree, that Braden-Harder teaches scoring metrics to take into account the frequency of a logical form triple occurring in a document (tf) which corresponds to the claimed keyfact frequency. *Id.* at 48 (citing Ex. 1020, 17:60–18:7; Ex. 1003 ¶ 229). Petitioner further provides sufficient evidence that to the extent "frequency of said keyfacts" requires using df to calculate idf, Messerly teaches using df in combination with tf which was a "well-known approach[] to ranking documents by relevancy." *Id.* at 48–49 (citing Ex. 1020, 2:42–46, 17:60–18:7; Ex. 1025, 12:33–36, 12:51–59, 13:8–10, 14:8–12; Ex. 1003 ¶¶ 232, 234, 235).

For the "generating a keyfact list of said document collection for a keyfact index structure" limitation, Petitioner further contends, and we agree, that, as an example, Braden-Harder describes an indexing engine that generates and stores a set of logical form triples for each indexed document in a Master Dataset 1030 (index structure). *Id.* at 50–51 (citing Ex. 1020, 7:12–16, 15:60–63, 16:61–62, 20:10–13, 20:50–61, 20:64–21:3, 21:7–25, Fig. 10A; Ex. 1003 ¶ 241). Petitioner further provides a detailed showing as to how the limitation is met if the claimed "keyfact index structure" requires

each of the specific tables in Figure 4 of the '908 patent. *Id.* at 52–54. Other than arguing that "Petitioner errs by conflating distinct claim requirements in a manner that impermissibly reads out limitations" (PO Resp. 14–19), which we address above, Patent Owner does not dispute Petitioner's specific showing with respect to the above limitation. *See generally id.*

> *d. Claim 6: "keyfact retrieving step for receiving said keyfact of said user query and said keyfacts of said document collection and defining a keyfact retrieval model in consideration of weigh factors according to a keyfact pattern and generating a retrieval result"*

For the above limitation, Petitioner contends, and we agree, that Braden-Harder teaches, during the retrieval process, obtaining a user's full-text query and generating logical form triples from the user query block and receiving logical form triples associated with the retrieved document records. Pet. 55–56 (citing Ex. 1020, 11:19–25, 12:16–29, 15:9–21, Fig. 12A; Ex. 1003 ¶ 257).

Petitioner further contends that Braden-Harder teaches "defining a keyfact retrieval model in consideration of weigh factors according to a keyfact pattern," as Braden-Harder teaches its retrieval model uses document weights calculated according to a logical-form-triple pattern. *Id.* at 57–58 (citing Ex. 1020, Fig. 8B). In particular, Petitioner contends, and we agree, that Braden-Harder teaches calculating a document weight by "weighting of matching logical forms . . . associated with a search query and each of the retrieved documents." *Id.* at 58 (quoting Ex. 1020, 5:8–11) (alteration in original). Petitioner further contends that the weights are based on the semantic relationships in the logical form triples, such as the words' functional roles, semantic roles, or semantic labels. *Id.* (citing Ex. 1020,

Tables 2–4, 5:41–44 ("different relation type, i.e., such as deep subject, deep object, operator and the like, that can occur in a logical form is assigned a predefined weight"), 16:21–25, 16:27–29 ("[t]he weight reflects a relative importance ascribed to that relation in indicating a correct semantic match between a query and a document"), Fig. 8A). Petitioner asserts, therefore, and we agree, that "Braden-Harder defines a retrieval model in consideration of weight constants according to logical form triples, like the '908 patent." *Id.* at 58–59 (citing Ex. 1003 ¶ 265). Petitioner further contends that to the extent the claimed "keyfact retrieval model" requires the use of a vector-space retrieval model, "then Braden-Harder in view of Grossman renders this obvious" because Braden-Harder acknowledges that use of a statistical vector-space model was conventional. *Id.* at 59 (citing Ex. 1020, 25:35–36). Petitioner asserts that Grossman teaches a vector-space model and further provides reasons for combining Braden-Harder with Grossman. *Id.* at 59–61 (citing Ex. 1010, 13, 16, 18, 80; Ex. 1020, 11:4–7, 17:16–18; Ex. 1003 ¶¶ 267, 272–275). Lastly, Petitioner contends that Braden-Harder (with Messerly) teaches using *tf-idf* weights augmented by its logic-form-triples weight constant that when implemented with Grossman calculates a similarity coefficient using vector-space mode that results in an order of retrieval results. *Id.* at 61–62 (citing *id.* § VI.A.1.c (1), § VI.A.1.d; Ex. 1020, 8:2–6, 16:35–40; Ex. 1003 ¶¶ 279–281).

Patent Owner argues that Braden-Harder's "predetermined" weights do not teach or suggest the claimed "weigh factors according to a keyfact pattern," which are described in the '908 patent as keyfact weight constants "determined experimentally on the basis of distribution of keyfact pattern of document collection." PO Resp. 24–28. Petitioner argues that we "considered and rejected Uniloc's previous identical arguments" in the

Decision to Institute and that "Uniloc adds nothing to dispute" our preliminary determinations. Pet. Reply 27.

We agree with Petitioner that Patent Owner's arguments directed to the "keyfact retrieving step" are the same arguments Patent Owner raised in the Preliminary Response. *Compare* PO Resp. 24–28, *with* Prelim. Resp. 27–31. For the same reasons discussed in the Decision to Institute, those arguments are not persuasive. *See* Dec. 33–34. That is, we do not agree that the "weigh factors" must be determined experimentally or cannot be predetermined weights as Patent Owner asserts. Moreover, the '908 patent describes determining the weight constants "on the basis of distribution of keyfact pattern of document collection" (Ex. 1001, 7:62–64) like Braden-Harder, which similarly describes that its weights "reflect[] a relative importance ascribed to that relation in indicating a correct semantic match between a query and a document." Ex. 1020, 16:27–29. In other words, Braden-Harder's weights are assigned based on the keyfact pattern, like in the '908 patent. Ex. 1003 ¶ 265.

e. Claim 6: Motivation and Rationale to Combine

Petitioner argues a person of ordinary skill in the art would have been motivated and found it obvious to combine the teachings of Braden-Harder and Grossman. *See, e.g.*, Pet. 52–53. Petitioner asserts that Grossman teaches that tables (or their equivalents), such as those shown in the table of Figure 4 of the '908 patent[21], were known for implementing document-index structures such as the indexed database in Braden-Harder. *Id.* (citing Ex. 1003 ¶ 244). Petitioner contends, and we agree, that "[s]uch a technique

[21] Petitioner makes this assertion assuming that the "keyfact index structure" is in the form of each of the specific tables in Figure 4 of the '908 patent. *See* Pet. 52–53.

for generating the index structure would have been an obvious implementation to a POSITA at the time of the '908 patent, as taught in Grossman." *Id.* at 53 (citing Ex. 1010, 134–137). Patent Owner does not dispute Petitioner's showing regarding the reasons to combine Braden-Harder and Grossman. *See generally* PO Resp.

For all of the reasons above, we determine that Petitioner has demonstrated by a preponderance of the evidence that claim 6 would have been obvious over Braden-Harder in view of Grossman.

f. Dependent Claims 7–12

Each of claims 7–12 depend directly or indirectly from independent claim 6. Petitioner identifies supporting portions of Braden-Harder and Grossman, along with arguments and supporting evidence as to how the combined teachings of Braden-Harder and Grossman account for the additional limitations of these dependent claims. Pet. 63–84.

For example, claim 7 depends from claim 6 and recites, "wherein said step of keyfact extracting comprises the steps of analyzing morphology of an input sentence and obtaining tag sequences of part-of-speech by attaching part-of-speech tags." Petitioner references its showing for claim 6 and argues, and we agree, that Braden-Harder analyzes morphology of input sentence to parse text into its constituent words using a stored lexicon. *Id.* at 63 (citing Ex. 1020, 5:33–35, 11:62–12:1, 12:30–34). Petitioner further argues, and we agree, that during its morphological analysis, "Braden-Harder (as explained in Heidorn) attaches each parsed word with a part-of-speech tag to obtain tag sequences." *Id.* (citing Ex. 1020, 12:41–42, 43; Ex. 1022, 3:36–40; Ex. 1003 ¶¶ 286–287). For the "selecting a tag sequence of part-of-speech out of said tag sequences of part-of-speech" limitation of claim 7, Petitioner again refers to its showing for claim 6 and argues, and we

agree, that Braden-Harder (with Heidorn) uses syntactic analysis to select a final tag sequence out of a set of tag sequences. *Id.* at 63–64 (citing Ex. 1022, 4:20–22, 4:56–58; Ex. 1003 ¶¶ 289–292). Claim 7 also recites, "extracting a keyfact pattern by applying said tag sequence of part-of-speech to a keyfact pattern rule." Petitioner refers to its showing for claim 6, and persuasively argues, with supporting testimony, that Braden-Harder uses semantic analysis to extract a logical form graph by applying its syntactic-parse tree (with the associated part-of-speech tag sequence) to at least one semantic rule. *Id.* at 64 (citing Ex. 1003 ¶¶ 294–295). Lastly, claim 7 recites, "applying said keyfact pattern to a keyfact pattern generation rule and generating a keyfact list." Petitioner refers to its showing for claim 6, and persuasively argues, with supporting testimony, that Braden-Harder applies the logical form graph to a series of graph-walk rules to generate a list of logical form triples. *Id.* at 65 (citing Ex. 1003 ¶¶ 212–218, 296–298).

Patent Owner argues that the "Petition is keyed to incorrect claim constructions which improperly attempt to incorporate limitations from dependent claims into independent claim 6." PO Resp. 29. We have addressed that argument in our claim construction discussion above. Patent Owner does not otherwise address separately Petitioner's arguments and supporting evidence as to how the combined teachings of Braden-Harder and Grossman account for the additional limitations of dependent claims 7–12. *See Id.* at 28–29; Pet. Reply 30; Sur-reply 1–22. We have reviewed Petitioner's arguments and supporting evidence as to how the combined teachings of Braden-Harder and Grossman accounts for the additional limitations of dependent claims 7–12, and we agree with and adopt Petitioner's analysis. *See* Pet. 63–84. Accordingly, Petitioner has demonstrated by a preponderance of the evidence that the subject matter of

dependent claims 7–12 would have been obvious over the combined teachings of Braden-Harder and Grossman.

E. Remaining Grounds Challenging Claims 6–12

For the reasons discussed above, Petitioner has shown, by a preponderance of the evidence that claims 6–12 are unpatentable as obvious over Braden-Harder and Grossman. In addressing this ground, we have addressed all of the challenged claims. *See* 35 U.S.C. § 318(a) (requiring the Board to "issue a final written decision with respect to the patentability of any patent claim challenged by the petitioner and any new claim added under section 316(d)"); *see also SAS Inst. Inc. v. Iancu*, 138 S. Ct. 1348, 1359 (2108) (holding that a petitioner "is entitled to a final written decision addressing all of the claims it has challenged"). Accordingly, we need not and do not decide whether Petitioner has shown by a preponderance of the evidence that claims 6–12 are unpatentable based on the remaining challenges. *Cf. In re Gleave*, 560 F.3d 1331, 1338 (Fed. Cir. 2009) (not reaching other grounds of unpatentability after affirming the anticipation ground); *see also Beloit Corp. v. Valmet Oy*, 742 F.2d 1421, 1423 (Fed. Cir. 1984) (holding that once a dispositive issue is decided, there is no need to decide other issues).

III. CONCLUSION

For the foregoing reasons, we determine that Petitioner has shown by a preponderance of the evidence that claims 6–12 of the '908 patent are unpatentable, as summarized in the following table:

Claims	35 U.S.C. §	References	Claims Shown Unpatentable	Claims Not Shown Unpatentable
6–12	103(a)[22]	Braden-Harder, Grossman	6–12	
6–12	103(a)	Braden-Harder, Heidorn, Messerly, Grossman		
7–9	103(a)	Braden-Harder, Grossman, Kucera		
7–9	103(a)	Braden-Harder, Heidorn, Messerly, Grossman, Kucera		
9	103(a)	Braden-Harder, Grossman, Miller		
9	103(a)	Braden-Harder, Heidorn, Messerly, Grossman, Miller		
9		Braden-Harder, Grossman, Kucera, Miller		
9		Braden-Harder, Heidorn, Messerly, Grossman, Kucera, Miller		

[22] As explained immediately above, we need not and do not decide whether Petitioner has shown by a preponderance of the evidence that claims 6–12 also would have been obvious based on the remaining grounds not addressed in this Decision.

Claims	35 U.S.C. §	References	Claims Shown Unpatentable	Claims Not Shown Unpatentable
Overall Outcome			6–12	

IV. ORDER

Accordingly, it is:

ORDERED that claims 6–12 of the '908 patent have been shown to be unpatentable; and

FURTHER ORDERED that, because this is a Final Written Decision, parties to the proceeding seeking judicial review of the Decision must comply with the notice and service requirements of 37 C.F.R. § 90.2.

IPR2020-00755
Patent 6,366,908 B1

For PETITIONER:

Erika Arner
Joshua Goldberg
FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP
erika.arner@finnegan.com
joshua.goldberg@finnegan.com

For PATENT OWNER:

Brett Mangrum
Ryan Loveless
James Etheridge
Brian Koide
Jeffrey Huang
ETHERIDGE LAW GROUP
brett@etheridgelaw.com
ryan@etheridgelaw.com
jim@etheridgelaw.com
brian@etheridgelaw.com
jeff@etheridgelaw.com

US006366908B1

(12) **United States Patent**
Chong et al.

(10) **Patent No.: US 6,366,908 B1**
(45) **Date of Patent: Apr. 2, 2002**

(54) **KEYFACT-BASED TEXT RETRIEVAL SYSTEM, KEYFACT-BASED TEXT INDEX METHOD, AND RETRIEVAL METHOD**

(75) Inventors: **Kyung Taek Chong; Myung-Gil Jang; MiSeon Jun; Se Young Park**, all of Taejon (KR)

(73) Assignee: **Electronics and Telecommunications Research Institute**, Taejon (KR)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **09/475,743**

(22) Filed: **Dec. 30, 1999**

(30) **Foreign Application Priority Data**

Jun. 28, 1999 (KR) 99-25035

(51) **Int. Cl.**[7] **G06F 17/30**

(52) **U.S. Cl.** **707/3**; 707/5; 707/6; 707/101; 707/201; 704/7; 704/9; 704/10; 382/177; 382/306

(58) **Field of Search** 707/1, 2, 3, 4, 707/5, 6, 102, 103, 104.1; 706/12, 45, 47; 704/7, 9, 10, 270.1

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,839,853 A		6/1989	Deerwester et al.	364/900
5,289,375 A	*	2/1994	Fukumochi et al.	364/419.02
5,535,382 A		7/1996	Ogawa	395/600
5,541,836 A	*	7/1996	Church et al.	364/419.07
5,598,557 A	*	1/1997	Doner et al.	395/605
5,708,829 A		1/1998	Kadashevich et al.	395/793
5,721,902 A	*	2/1998	Schultz	395/604
5,771,378 A	*	6/1998	Holt et al.	395/605

OTHER PUBLICATIONS

Jang, Ho Wook and Se Young Park, "Keyfact Concept For An Information Retrieval System," Proceedings of Natural Language Proc. Pacific Rim Symposium, pp. 510–513, 1995.

Jun, MiSeon and Se Young Park, "Keyfact–Based Information Retrieval System," International Symposium on Digital Library, pp. 521–524, 1997.

Arampatzis et al., "Phrase–Based Information Retrieval," *Journal of Information Proc. & Management 34*(6):1–19, 1998.

Robert Krovetz and W. Bruce Croft:"Lexical ambiguity and information retrieval", Apr. 1992, ACM, vol. 10, pp. 115–141.*

Michael Sussna: "Word sense disambiguation for free–text indexing using a massive semantic network", 1993, ACM, pp. 67–74.*

* cited by examiner

Primary Examiner—Thomas Black
Assistant Examiner—Jacques Veillard
(74) *Attorney, Agent, or Firm*—E. Russell Tarleton; SEED IP Law Group PLLC

(57) **ABSTRACT**

A keyfact-based text retrieval method and a keyfact-based text index method that describes the formalized concept of a document by a pair comprising an object that is the head and a property that is the modifier and uses the information described by the pairs as index information for efficient document retrieval. A keyfact-based text retrieval system includes keyfact extracting, keyfact indexing, and keyfact retrieving. The keyfact extracting analyzes a document collection and a query and extracts keywords and keyfacts. The keywords do not have part-of-speech ambiguity and the keyfacts are extracted from the keywords. The keyfact indexing calculates the frequency of the keyfacts and generates a keyfact list of the document collection for a keyfact index structure. The keyfact retrieving receive a keyfact of the query and keyfacts of the document collection and defines a keyfact-based retrieval model in consideration of a weight factor of the keyfact pattern and generates a retrieval result. The retrieval result is a document similar to the query.

12 Claims, 5 Drawing Sheets

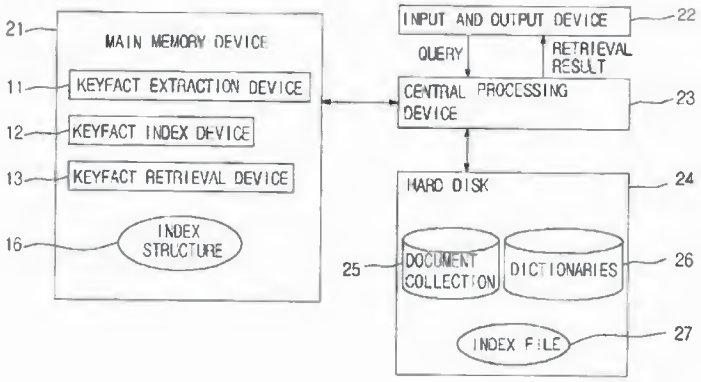


FIG. 1

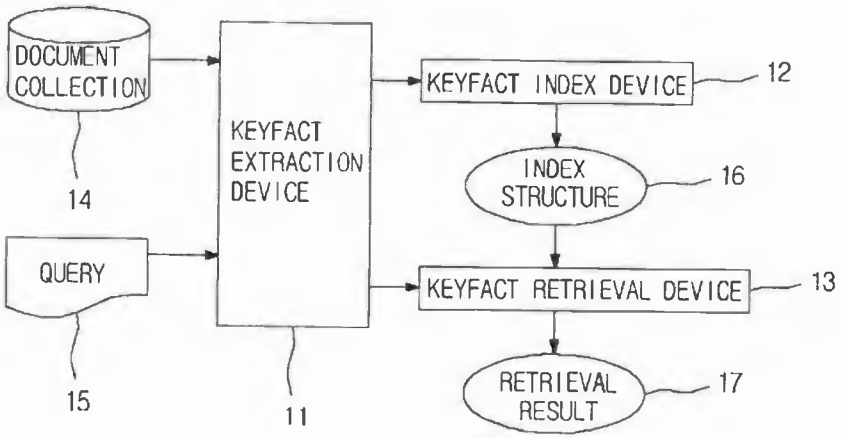

DOCUMENT COLLECTION
14
QUERY
15
KEYFACT EXTRACTION DEVICE
11
KEYFACT INDEX DEVICE
12
INDEX STRUCTURE
16
KEYFACT RETRIEVAL DEVICE
13
RETRIEVAL RESULT
17

FIG. 2

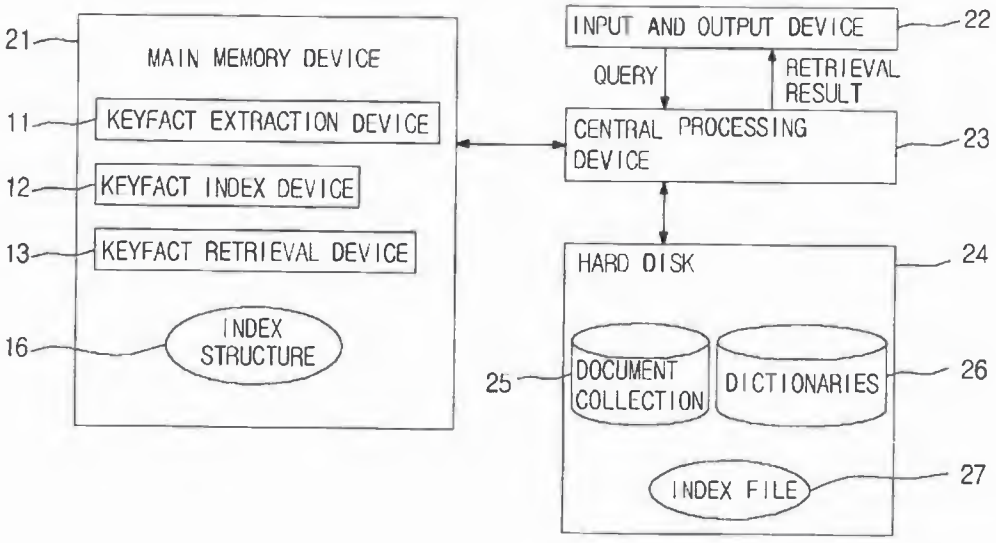

21
MAIN MEMORY DEVICE
11
KEYFACT EXTRACTION DEVICE
12
KEYFACT INDEX DEVICE
13
KEYFACT RETRIEVAL DEVICE
16
INDEX STRUCTURE
INPUT AND OUTPUT DEVICE
22
QUERY
RETRIEVAL RESULT
CENTRAL PROCESSING DEVICE
23
HARD DISK
24
25
DOCUMENT COLLECTION
DICTIONARIES
26
INDEX FILE
27

FIG. 3

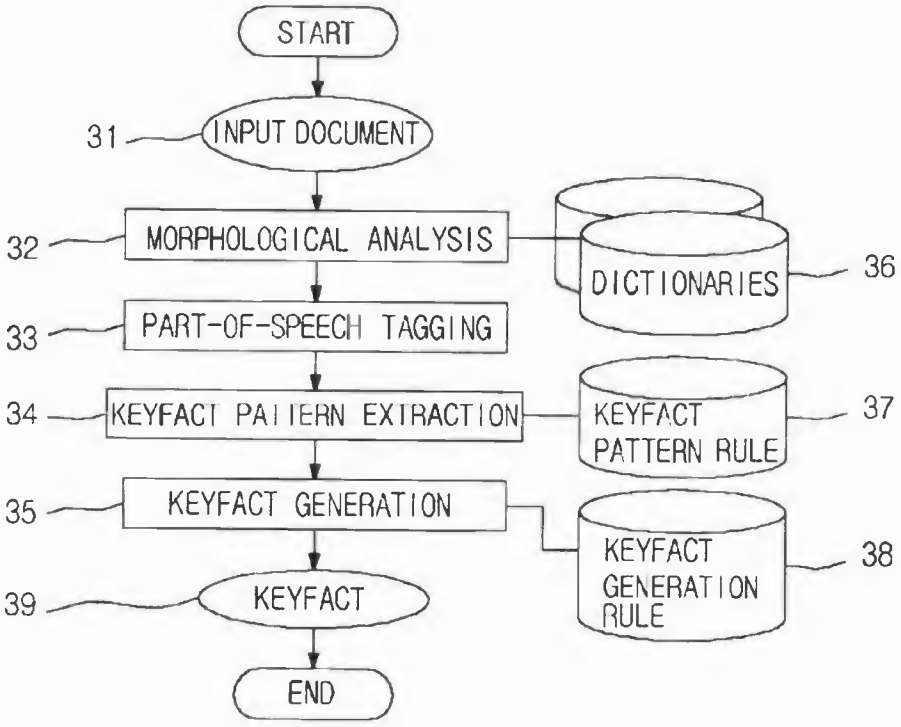

START
31
INPUT DOCUMENT
32
MORPHOLOGICAL ANALYSIS
DICTIONARIES
36
33
PART-OF-SPEECH TAGGING
34
KEYFACT PATTERN EXTRACTION
KEYFACT PATTERN RULE
37
35
KEYFACT GENERATION
KEYFACT GENERATION RULE
38
39
KEYFACT
END

FIG. 4

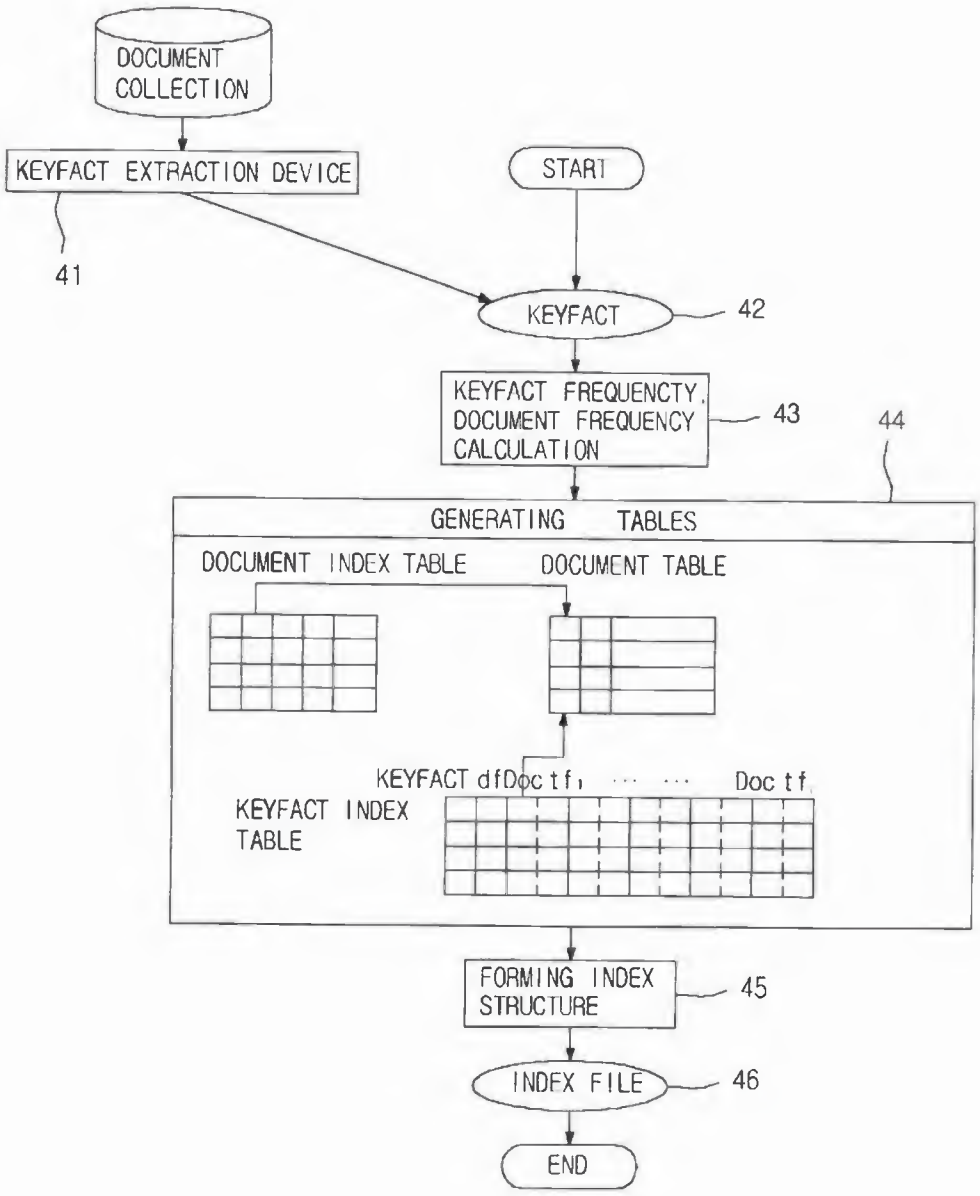

DOCUMENT COLLECTION
KEYFACT EXTRACTION DEVICE
41
START
KEYFACT
42
KEYFACT FREQUENCTY, DOCUMENT FREQUENCY CALCULATION
43
44
GENERATING TABLES
DOCUMENT INDEX TABLE
DOCUMENT TABLE
KEYFACT dfDoc t f1 ··· ··· Doc t f
KEYFACT INDEX TABLE
FORMING INDEX STRUCTURE
45
INDEX FILE
46
END

FIG. 5

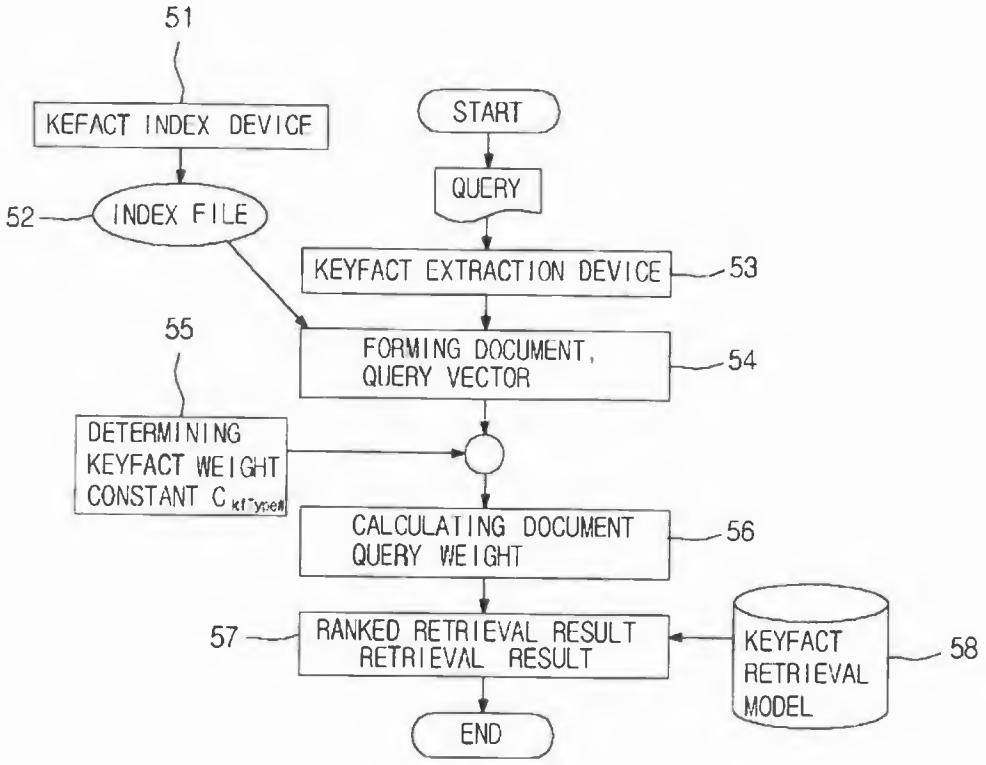
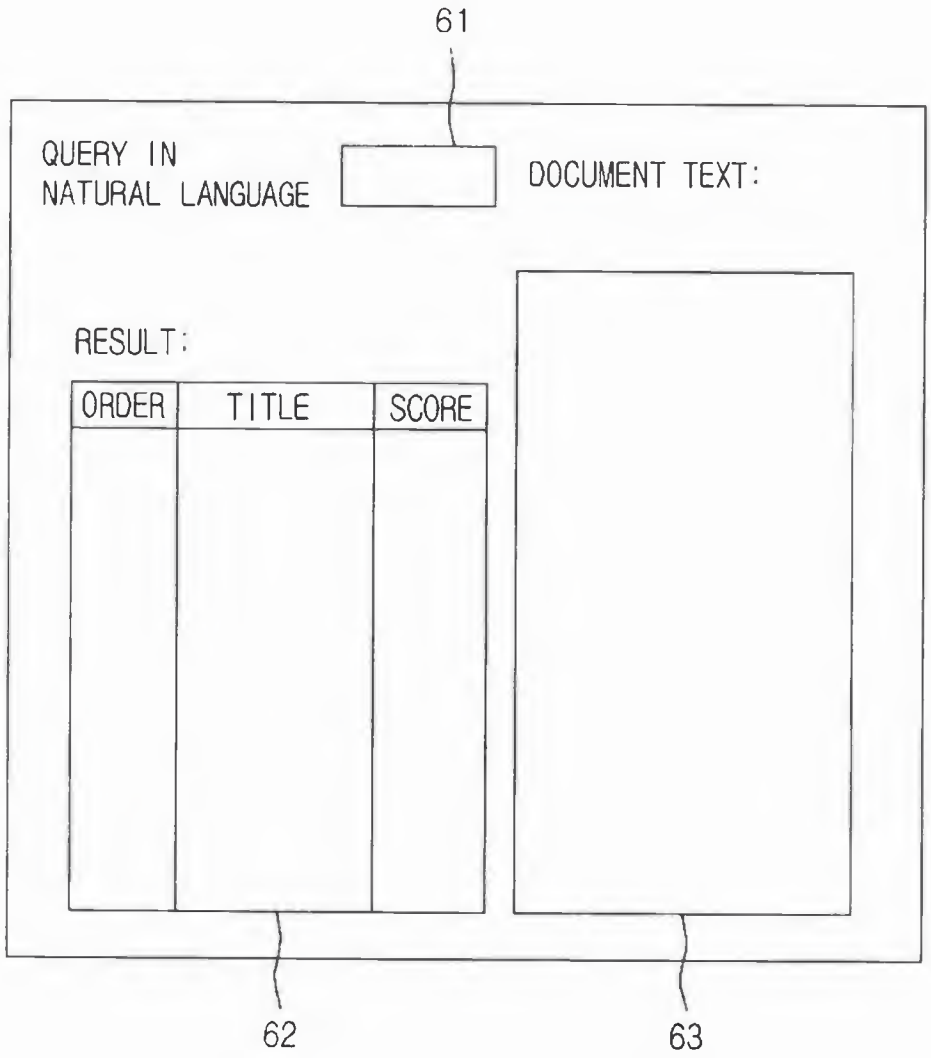

51
KEFACT INDEX DEVICE
START
QUERY
52
INDEX FILE
KEYFACT EXTRACTION DEVICE
53
55
FORMING DOCUMENT, QUERY VECTOR
54
DETERMINING KEYFACT WEIGHT CONSTANT $C_{kf\,type}$
CALCULATING DOCUMENT QUERY WEIGHT
56
57
RANKED RETRIEVAL RESULT RETRIEVAL RESULT
KEYFACT RETRIEVAL MODEL
58
END

FIG. 6


61
QUERY IN
NATURAL LANGUAGE
DOCUMENT TEXT:
RESULT:
ORDER
TITLE
SCORE
62
63

KEYFACT-BASED TEXT RETRIEVAL SYSTEM, KEYFACT-BASED TEXT INDEX METHOD, AND RETRIEVAL METHOD

TECHNICAL FIELD

The present invention relates to a keyfact-based text retrieval method and a keyfact-based text index method. In particular, the methods describe the formalized concept of a document as a pair comprising an object that is the head and a property that is the modifier, and uses the information described by the pair as index information for efficient document retrieval.

BACKGROUND OF THE INVENTION

A keyfact means an important fact contained in sentences which constitute a document. The keyfact is represented by an object and property information through syntactic analysis of the sentence.

The keyword-based text retrieval method was the main stream in conventional text retrieval methods. However, the precision of the keyword-based text retrieval method was not good due to the following reasons. First, the meaning of the document is not precisely represented and the representativeness of document expression is low because the document is represented by keywords, which are nouns. This is a fundamental reason for poor retrieval precision. Second, when a query includes a natural language phrase or a natural language sentence or keywords, the intention of the user's query is not reflected precisely in a keyword-based text retrieval method because the query is expressed by keywords. Therefore, the keyword-based text retrieval method has a fundamental limitation in retrieval precision because it performs document retrieval by keywords. As a result, because the keyword-based text retrieval system provides such low level of retrieval precision, it causes a number of unnecessary retrievals and therefore precious resources, such as time and effort, are wasted.

Recently, a number of studies have been performed in the area of phrase-based text retrieval methods in order to compromise such defects of the keyword-based retrieval method. The phrase-based text retrieval methods extract a precise phrase pattern through a morphological-syntactic normalization process and perform indexing and retrieval by extracted phrase. Therefore, the phrase-based retrieval method performs more precise text retrieval than the keyword-based text retrieval method but performs less precise text retrieval than a concept-based text retrieval method, which expresses text by concept units.

A new approach to keyfact-based text retrieval methods has been proposed in order to overcome the shortcomings of the keyword-based text retrieval method and generalize phrase-based text retrieval method. In the keyfact-base text retrieval method, a part of text that represent the same meaning is described as a keyfact. Since the keyfact-based retrieval method is a sort of concept-based retrieval method, and therefore indexing and retrieval of the keyfact-based retrieval method are performed with the unit of the keyfact, precision of the retrieval is greatly improved.

In the keyfact-based retrieval method, it is desirable that phrases or words having the same meaning are indexed as the same indexing terms. For example, noun phrases including "the retrieval of information" as a subset of "the efficient retrieval of information", "the retrieval of the distributed information", and "the fast retrieval of the distributed information" must have common indices which can be possibly generated from "the retrieval of information" as subsets and

recognize also them as different meaning with subtle conceptual different indexes at the same time.

Since the keyword-based retrieval method doesn't recognize the conceptual difference between "the retrieval of the information" and "the efficient retrieval of the information", users are not able to retrieve the exact document that is desired.

SUMMARY OF THE INVENTION

A keyfact-based retrieval method, which extracts the precise keyfact pattern using the natural language processing techniques and indexes documents with the unit of the keyfact, is provided.

In addition, a keyfact-based retrieval method, which extracts precise keyfact patterns included in a natural query of a user using the natural language processing techniques and retrieves documents similar to the query in the keyfact-based index file, is provided.

In addition, a keyfact-based retrieval method, which retrieves and indexes documents with the unit of keyfact, is provided.

A keyfact-based text retrieval system of the present invention includes keyfact extracting means, keyfact indexing means, and keyfact retrieving means. The keyfact extracting means analyze a document collection and a user query, and extracting keywords not having part-of-speech ambiguity from the document collection and the user query, and respectively extracting keyfacts of the document collection and the user query from the keywords. The keyfact indexing means for calculating the frequency of the keyfacts of the document collection and generating a keyfact list of the document collection for a keyfact index structure. The keyfact retrieving means for receiving the keyfact of the user query and the keyfacts of the document collection and defining a keyfact retrieval model in consideration of weight factors according to a keyfact pattern and generating a retrieval result.

The keyfact extracting means includes morphology analyzing means, part-of-speech tagging means, keyfact pattern extracting means, and keyfact generating means. The morphology analyzing means analyze morphology of an input sentence and obtaining tag sequences of part-of-speech by attaching part-of-speech tags. The part-of-speech tagging means selects a tag sequence of part-of-speech out of the tag sequences of part-of-speech. The tag sequence of part-of-speech is precise. The keyfact pattern extracting means extracts a keyfact pattern by applying the tag sequences of part-of-speech to a keyfact pattern rule. The keyfact generating means applies the keyfact pattern to a keyfact pattern generation rule and generating a keyfact list, which is a set of keyfact terms.

The keyfact indexing means includes frequency calculating means, table generating means, and keyfact indexing means. The frequency calculating means calculates a frequency of various keyfacts and a document frequency of the keyfacts. The various keyfacts are included in the document collection, and the document frequency is the number of documents contained the various keyfacts. The table generating means generates a document index table, a document table, and a keyfact index table of the document collection. The keyfact indexing means forms a keyfact index structure. The keyfact index structure has information regarding document frequency, document identifier, and keyfact frequency in each corresponded documents.

The keyfact retrieving means includes following means. A means forms a document and a user query vector with an

index file and the keyfact of the user query. The index file generated by the keyfact indexing means. The keyfact of the user query generated by the keyfact extracting means. A means determines keyfact weight constants in accordance with the keyfact pattern. A means calculates keyfact weights for the document and the user query by applying the keyfact weight constants to the document and the user query vector. The retrieval results displaying means displays the retrieval result by applying the keyfact weights to keyfact retrieval model. The retrieval result indicates documents with a keyfact similar to the keyfact of the user query.

A keyfact-based text retrieving method of the present invention includes keyfact extracting step, keyfact indexing step, and keyfact retrieving step. The keyfact extracting step is to analyze a document collection and a user query, and extracts keywords without part-of-speech ambiguity from the document collection and the user query, and respectively extracts keyfacts of the document collection and the user query from the keywords. The keyfact indexing step is to calculates the frequency of the keyfacts of the document collection and generates a keyfact list of the document collection for a keyfact index structure. The keyfact retrieving step is to receives the keyfact of the user query and the keyfacts of the document collection and defines a keyfact retrieval model in consideration of weigh factors according to the keyfact pattern and generates the retrieval result.

The step of keyfact extracting includes the following steps. The first step is to analyze morphology of an input sentence and obtaining tag sequences of part-of-speech by attaching part-of-speech tags. The second step is to select a tag sequence of part-of-speech out of the tag sequences of part-of-speech. The third step is to extract a keyfact pattern by applying the tag sequence of part-of-speech to a keyfact pattern rule. The fourth step is to apply the keyfact pattern to a keyfact pattern generation rule and generating a keyfact list.

The step of analyzing morphology includes the following steps. The first step is to divide the input sentence into words. The second step is to perform morphological analysis on the words using part-of-speech dictionaries. The third step is to perform morphological variation and recover prototypes. The fourth step is to obtain the tag sequence of part-of-speech by tagging part-of-speech tags in accordance with the result of the morphological analysis.

The part-of-speech dictionaries include a noun dictionary, a verb dictionary, an adjective dictionary, an adverb dictionary, a preposition dictionary, a conjunction dictionary and a stop-word lexicon.

The step of keyfact indexing includes the following steps. The first step is to calculate a frequency of various keyfacts and a document frequency of the keyfact. The second step is to generate a document index table, a document table and a keyfact index table of the document collection. The third step is to form a keyfact index structure including document frequency, document identifier and keyfact frequency.

The step of keyfact retrieving includes the following steps. The first step is to form a document and a user query vector with an index file and a keyfact of the user query. The second step is to determine keyfact weight constants in accordance with the keyfact pattern. The third step is to calculate keyfact weights for the document and the user query by applying the keyfact weight constants to the document and the user query vector. The fourth step is to display the retrieval result by applying the keyfact weights to the keyfact retrieval model. The retrieval result indicates documents with a keyfact similar to the keyfact of the user query.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. **1** is a block diagram illustrating a keyfact-based text retrieval system of the present invention;

FIG. **2** is a block diagram illustrating a hardware structure of a keyfact-based text retrieval system in accordance with an embodiment of the present invention;

FIG. **3** is a block diagram illustrating a keyfact extraction device of a keyfact-based text retrieval system in accordance with an embodiment of the present invention;

FIG. **4** is a block diagram illustrating a keyfact index device of a keyfact-based text retrieval system in accordance with an embodiment of the present invention;

FIG. **5** is a block diagram illustrating a keyfact retrieval device of a keyfact-based text retrieval system in accordance with an embodiment of the present invention; and

FIG. **6** is a screen image illustrating a document retrieval result in response to a query.

DETAILED DESCRIPTION OF THE INVENTION

FIG. **1** is a block diagram illustrating a keyfact-based text retrieval system of the present invention. The keyfact-based text retrieval system comprises a keyfact extraction device **11**, a keyfact index device **12**, and a keyfact retrieval device **13**. FIG. **2** is a block diagram illustrating a hardware structure of a keyfact-based text retrieval system in accordance with an embodiment of the present invention.

As shown in FIG. **2**, the main memory device **21** includes a keyfact extraction device, a keyfact index device **12**, a keyfact retrieval device **13**, and an index structure **16**. The central processing device **23** supervises the keyfact-based text retrieval. A hard disk **24** stores document collection **25**, dictionaries for keyfact retrieval **26**, and an index file that is the result of the keyfact index. The index file **27** is loaded onto the main memory as an index structure **16** and the keyfact retrieval device **13** uses the index file. The input and output device **22** receives a query from a user and generates retrieval results to the user.

Now, the keyfact-based text retrieval system in accordance with the present invention is explained with reference to FIG. **1**. Once a document collection **14** or a query **15** is given, the keyfact extraction device **11** extracts words without ambiguity by performing morphological analysis and tagging. The keyfact generation rule is applied to the words and then the keyfacts are extracted.

The keyfact index device **12** indexes the document collection **14** or the query with the unit of keyfact and calculates the frequencies of the keyfacts. The frequencies of the keyfacts are stored into the index structure **16** with the document ID information. The keyfact retrieval device **13** orders documents using the similarity calculation method and shows retrieval results. The similarity calculation method considers document collection and keyfact weights with the help of a keyfact-based text retrieval model. In a keyfact-based text retrieval, when a document collection **14** or a query is given, the keyfact extraction device **11** expresses it in the unit of keyfacts. All keyfacts express semantic relation between words in the form of [object, property]. Keyfacts can be categorized by configurations of an object and a property. Parts of text that express the same conceptual meaning in the document collection or the query are categorized into the same keyfact type. The keyfact extraction device will be reviewed in detail below with FIG. **3**.

The keyfact index device **12** indexes the extracted keyfacts with frequency information. In other words, the keyfact

index device **12** calculates frequencies of the various forms of keyfacts included in the documents and generates a keyfact list of the document collection. Therefore, an index structure **16** that reflects keyfacts is created and the index file is stored. The keyfact index device **12** will be reviewed in detail below with FIG. **4**.

When the keyfact retrieval device **13** receives a query, it retrieves appropriate documents on the basis of the keyfact-based retrieval method. The keyfact retrieval model is defined by considering weights of keyfact patterns. The similarity between the query and the documents is calculated and appropriate documents for the query are shown as a result in the order of the similarity. The keyfact retrieval device **13** will be reviewed in detail below with FIG. **5**.

As shown in FIG. **3**, the keyfact extraction device **11** analyzes a document and generates keyfacts through the processes of morphological analysis, part-of-speech tagging, keyfact pattern extraction, and keyfact generation.

A document is supplied at stage **31** and morphological analysis is performed at stage **32**. A sentence in the document is divided into words and the morphological analysis is performed with dictionaries **36** at stage **32**. The morphological variation is considered in order to recover prototypes. The dictionaries **36** include a noun dictionary, a verb dictionary, an adjective dictionary, an adverb dictionary, a preposition dictionary, a conjunction dictionary, and a stop-word lexicon. In some cases, a part-of-speech of a word is determined by rules without dictionaries.

The part-of-speech tag in dictionaries **36** includes noun (N), verb (V), adjective (A), preposition (P), and stop-word (S). The noun is further divided into proper noun (NQ), name noun (NN), vocative noun (NV), unit nouns (NJ), predicate noun (NU), non-predicate noun (NX), etc. The reason for such division is that the class of noun determines the object or the property of the keyfacts.

For example, in a sequence of words having two or three nouns in a row, it is likely that name noun (NN), proper noun (NQ), and non-predicate noun (NX) are objects and vocative noun (NV), unit noun Id), and predicate noun (NU) are properties. Additionally, in a phrase having proper noun (NQ), name noun (NN), and non-predicate noun (NX), the order of priority of nouns in the object is name noun (NN)>proper noun (NQ)>non-predicate noun (NX).

The preposition is divided into the possessive preposition (PO) which is used as "of" and the positional preposition (PP) and etc. The adjective or the variated verb which makes up the noun is tagged as a pronoun (MP), which is a separate keyfact tag. For example, in analyzing "the fast retrieval of the distributed information" with morphological analysis, a result of the sequence of the tag would be "S (stop-word) A (adjective) NV (vocative noun) PO (possessive preposition) S (stop-word) V-ed (verb) NV (vocative noun). The V-ed (verb) is a modified form of verb and makes up the noun. Like the A (adjective), the V-ed (verb) is converted into a keyfact tag MP and the sequence of nouns is converted into a keyfact tag KEY. The final result would become "NMP KEY PO MP KEY".

Once the stage **32** of morphological analysis is performed, various results are obtained.

At stage of **33** in which part-of-speech tagging is performed, a precise sequence of tags is chosen among the various results of the morphological analysis. In other words, the part-of-speech tags obtained from the morphological analysis are used at the stage of part-of-speech tagging. The modified form of verb that makes up a noun or an adjective is converted into a modifier (MP) and the sequence of nouns is converted into KEY tag. The exemplary sentence "the fast retrieval of the distributed information" shows the final sequence of tags "MP KEY PO MP KEY".

Once the final sequence of tags in response to the input sentence is obtained, the stage of keyfact pattern extraction **34** searches the keyfact pattern rule **37** and extracts meaningful keyfact patterns necessary for keyfact generation. The keyfact pattern rule **37** which is used for keyfact pattern extraction describes keyfact patterns as to the sequence of the input tags. A part of the keyfact pattern rule is illustrated at following table 1.

TABLE 1

Keyfact pattern	Keyfact term list
KEY1 PO KEY2 *(the retrieval of information)*	[KEY2, KEY1], [KEY1, NIL], [KEY2, NIL], [KEY2 KEY1, NIL] *[information, retrieval]*, *[information, NIL]*, *[retrieval, NIL]*, *[information retrieval, NIL]*
KEY1 PO MP KEY2 *(the retrieval of the distributed information)*	[KEY2, KEY1], [KEY1, NIL], [KEY2, NIL], [KEY2 KEY1, NIL], [KEY2, MP]
MP KEY1 PO KEY2 *(the fast retrieval of information)*	[KEY2, KEY1], [KEY1, NIL], [KEY2, NIL], [KEY2 KEY1, NIL], [KEY1, MP]
MP1 KEY1 PO MP2 KEY2 *(the fast retrieval of the distributed information)*	[KEY2, KEY1], [KEY1, NIL], [KEY2, NIL], [KEY2 KEY1, NIL], [KEY1, MP1], [KEY2, MP2]

(Note:
The italic is the examples.)

The final sequence of tags "MP KEY PO MP KEY" obtained from "the fast retrieval of the distributed information" is applied to the keyfact pattern rule and the keyfact pattern "MP1 KEY1 PO MP2 KEY2" is the result.

Keyfact terms that have forms of [object, property] are generated as to the input keyfact pattern at the stage of the keyfact generation **35** by searching the keyfact generation rule **38**. The object is a noun or a compound noun represented by a keyword and the property is a verbal word or a noun that makes up another noun, or a prototype of a verbal word.

The keyfact generation rule includes possible keyfact lists, each of which can be generated in each keyfact pattern. In the example stated above, if the keyfact pattern "MP1 KEY1 JY MP2 KEY2" is applied to the keyfact generation stage, "[KEY2, KEY1], [KEY1, NIL], [KEY2, NIL], [KEY2 KEY1, NIL], [KEY1, MP1], [KEY2, MP2]" is going to be the outcome. That is, a keyfact list **39** "[information, retrieval], [retrieval, NIL], [information, NIL], [information retrieval, NIL], [retrieval, fast], [information, distributed]" is obtained from the keyfact pattern "the fast retrieval of the distributed information".

The keyfact index device is now reviewed in detail with FIG. **4**.

The keyfact index device calculates statistical frequencies of keyfacts in a document obtained from the keyfact extraction device **11** and forms the index structure. Therefore, index information is efficiently maintained and processed by the keyfact index device. Each index term of the keyfact index device is an extracted keyfact term representing each document.

For each document, the keyfact frequency (tf) and document frequency of the keyfact (df) are calculated in order to obtain the frequency information of the keyfacts.

Next, supplementary tables such as a document index table, a document table, and a keyfact index table are generated to form an efficient index structure **44**. The document index table contains keyfacts of the document, the frequency information. The document table includes a real document text. The keyfact index table is the main table that includes the document frequency (df) of each keyfact, and pair list of the document identifier of each keyfact and the frequency information within a document (tf).

Next, an index structure is formed in the unit of the keyfact and an index file is stored. Efficient storage structures like the B+ tree can be used for the index structure. The inverted file structure of the keyfact index table is used as posting information file structure.

A part of the result of the keyfact index is shown in the following table 2.

TABLE 2

Keyfact index	Df	Document id: frequency
[thorn, sharp]	2	(162:1)(197:1)
[thorn, dull]	3	(102:2)(188:3)(193:1)
...		
[reed, NIL]	2	(6:2)(29:1)
[reed field, NIL]	1	(6:1)
[branch, NIL]	4	(21:1)(33:2)(88:1)(90:3)
[Dahurian buckhorn family, NIL]	1	(102:1)

At table 2, in case of [branch, NIL], "branch" appears at 4 documents and therefore the document frequency (df) for keyfact index [branch, NIL] is four. In addition, "branch" appears once in document **21**, twice in document **33**, once in document **88**, and three times in document **90**.

The keyfact retrieval device **13** is now reviewed in detail with FIG. **5**. The keyfact retrieval device forms the document vector and query vector with the keyfact, which is supplied from the keyfact extraction device **53**, and the index file **52** generated by the keyfact index device **51**.

The keyfact weight constants ($C_{KfType\#}$), which are fit for the attribute of a document collection, are determined **55** before calculating the keyfact weights from document and query vector. Table **3** shows that keyfact weight constants are assigned to various patterns of keyfacts.

TABLE 3

Types	Keyfact pattern	Weight constants
Type 1	[KEY, NIL]	$C_{KfTypeI}$
Type 2	[KEY, MP] or [KEY, VH/VB]	$C_{KfTypeII}$
Type 3	[KEY1, KEY2]	$C_{KfTypeIII}$
Type 4	[KEY1 KEY2, NIL] or [KEY2 KEY 1, NIL]	$C_{KfTypeIV}$
Type 5	[KEY1 KEY2 KEY3]	$C_{KfTypeV}$
...	...	...

The keyfact weight constants are assigned with the sequence like $C_{KfTypeI}<C_{KfTypeII}<C_{KfTypeIII}<C_{KfTypeIV}<C_{KfTypeV}<\ldots$ and do important role for the precision of keyfact-based text retrieval. Therefore, weight constants are determined experimentally on the basis of distribution of keyfact pattern of document collection.

The keyfact weight constant is applied to the following equation 1 and the result of equation 1, a keyfact weight (W_{xk}), is used in the keyfact-based text retrieval model.

$$w_{xk} = tf_{xk} \cdot \log\left(\frac{N+1}{df_k}\right) \cdot C_{KfType\#} \qquad \text{[Equation 1]}$$

W_{xk}: a keyfact weight

tf_{xk}: frequency of a keyfact

N: size of a document

df_k: document frequency of a keyfact

$C_{kfType\#}$: a keyfact weight constant

Conventionally, only the frequency of keywords ($tf_{keyword}$), the document frequency of keywords ($df_{keyword}$), and the number of the documents in a document collection are considered in calculating the keyword weight in the keyword-based text retrieval system. However, the keyfact weight constant ($C_{kfType\#}$) of the keyfact pattern is also reflected in calculating keyfact weights in the keyfact-based retrieval system, so as to make it possible to index and retrieve in the unit of a keyfact.

Next, the similarity of the document appropriate for the query is calculated by employing the keyfact retrieval model based upon the vector space model. The result of the similarity calculation determines the order of appropriate documents **57**.

FIG. **6** shows a screen image for illustrating a document retrieval result in response to a query. A user makes a query in query section **61** with natural language. The keyfact is extracted by the keyfact-based text retrieval system and the documents close to the query are found. The result of the retrieval of the query is displayed at the document retrieval result screen **62** in the order of similarity. Document title and weight are also displayed with the order of similarity. In addition, if the document displayed is selected, document text screen **63** shows the contents of text of the document.

According to the present invention, texts of document collection and user queries are expressed, indexed and retrieved by concept-based keyfacts. Therefore, more precise retrieval results are achievable. Additionally, since indexing and retrieval with high precision are possible, time and efforts can be minimized, the keyfact-based retrieval method in accordance with the present invention can be used in various applications. Especially, digital library, text and annotation based multimedia information retrieval of broadcasting station, internet application, information retrieval of electronics commercial trading, and education/medical/military application areas can take advantage of the present invention.

Although representative embodiments of the present invention have been disclosed for illustrative purposes, those skilled in the art will appreciate that various modifications, additions and substitutions are possible, without departing from the scope and spirit of the invention as recited in the accompanying claims.

What we claim:

1. A keyfact-based text retrieval system comprising:

keyfact extracting means for analyzing a document collection and a user query, and extracting keywords not having part-of-speech ambiguity from said document collection and said user query, and respectively extracting keyfacts of said document collection and said user query from said keywords;

keyfact indexing means for calculating the frequency of said keyfacts of said document collection and generating a keyfact list of said document collection for a keyfact index structure; and

keyfact retrieving means for receiving said keyfact of said user query and said keyfacts of said document collec-

tion and defining a keyfact retrieval model in consideration of weight factors according to a keyfact pattern and generating a retrieval result.

2. The keyfact-based text retrieval system of claim 1, wherein said keyfact extracting means comprises:

morphology analyzing means for analyzing morphology of an input sentence and obtaining tag sequences of part-of-speech by attaching part-of-speech tags;

part-of-speech tagging means for selecting a tag sequence of part-of-speech out of said tag sequences of part-of-speech, said tag sequence of part-of-speech being precise;

keyfact pattern extracting means for extracting a keyfact pattern by applying said tag sequences of part-of-speech to a keyfact pattern rule; and

keyfact generating means for applying said keyfact pattern to a keyfact pattern generation rule and generating a keyfact list.

3. The keyfact-based text retrieval system of claim 1, wherein said keyfact indexing means comprises:

frequency calculating means for calculating a frequency of various keyfacts and a document frequency of said keyfacts, said various keyfacts being included in said document collection, said document frequency being the number of documents containing said various keyfacts;

table generating means for generating a document index table, a document table, and a keyfact index table of said document collection; and

keyfact indexing means for forming a keyfact index structure having information regarding document frequency, document identifier, and keyfact frequency.

4. The keyfact-based text retrieval system of claim 1, wherein said keyfact retrieving means comprises:

means for forming a document and a user query vector with an index file and said keyfact of said user query, said index file generated by said keyfact indexing means, said keyfact of said user query generated by said keyfact extracting means;

means for determining keyfact weight constants in accordance with said keyfact pattern;

means for calculating keyfact weights for said document and said user query by applying said keyfact weight constants to said document and said user query vector: and

results displaying means for displaying a retrieval result by applying said keyfact weights to keyfact retrieval model.

5. The keyfact-based text retrieval system of claim 4, wherein said retrieval result indicates documents with a keyfact similar to said keyfact of said user query.

6. A keyfact-based text retrieving method comprising:

keyfact extracting step for analyzing a document collection and a user query, and extracting keywords without part-of-speech ambiguity from said document collection and said user query, and respectively extracting keyfacts of said document collection and said user query from said keywords;

keyfact indexing step for calculating the frequency of said keyfacts of said document collection and generating a keyfact list of said document collection for a keyfact index structure; and

keyfact retrieving step for receiving said keyfact of said user query and said keyfacts of said document collection and defining a keyfact retrieval model in consideration of weigh factors according to a keyfact pattern and generating a retrieval result.

7. The keyfact-based text retrieving method of claim 6, wherein said step of keyfact extracting comprises the steps of:

analyzing morphology of an input sentence and obtaining tag sequences of part-of-speech by attaching part-of-speech tags;

selecting a tag sequence of part-of-speech out of said tag sequences of part-of-speech;

extracting a keyfact pattern by applying said tag sequence of part-of-speech to a keyfact pattern rule; and

applying said keyfact pattern to a keyfact pattern generation rule and generating a keyfact list.

8. The keyfact-based text retrieving method of claim 7, wherein said step of analyzing morphology comprises the steps of:

dividing said input sentence into words;

performing morphological analysis on said words using part-of-speech dictionaries;

performing morphological variation and recovering prototypes; and

obtaining said tag sequence of part-of-speech by tagging part-of-speech tags in accordance with the result of said morphological analysis.

9. The keyfact-based text retrieving method of claim 8, wherein said part-of-speech dictionaries comprise a noun dictionary, a verb dictionary, an adjective dictionary, an adverb dictionary, a preposition dictionary, a conjunction dictionary and a stop-word lexicon.

10. The keyfact-based text retrieving method of claim 6, wherein said step of keyfact indexing comprising the steps of:

calculating a frequency of various keyfacts and a document frequency of said keyfact;

generating a document index table, a document table and a keyfact index table of said document collection; and

forming a keyfact index structure including document frequency, document identifier and keyfact frequency.

11. The keyfact-based text retrieving method of claim 6, wherein said step of keyfact retrieving comprising the steps of:

forming a document and a user query vector with an index file and a keyfact of said user query;

determining keyfact weight constants in accordance with said keyfact pattern;

calculating keyfact weights for said document and said user query by applying said keyfact weight constants to said document and said user query vector; and

displaying a retrieval result by applying said keyfact weights to said keyfact retrieval model.

12. The keyfact-based text retrieving method of claim 11, wherein said retrieval result indicates documents with a keyfact similar to said keyfact of said user query.

* * * * *

CERTIFICATE OF COMPLIANCE
WITH TYPE-VOLUME LIMITATION, TYPEFACE REQUIREMENTS,
AND TYPE STYLE REQUIREMENTS

1. This brief complies with the type-volume limitation of Federal Circuit Rule 32(b).

The brief contains 8,930 words, excluding the parts of the brief exempted by Federal Rule of Appellate Procedure 32(f) and Fed. Cir. R. 32(b).

2. This brief complies with the typeface requirements of Federal Rule of Appellate Procedure 32(a)(5) and the type style requirements of Federal Rule of Appellate Procedure 32(a)(6).

The brief has been prepared in proportionally spaced typeface using MS Word – Office 365 in Times New Roman 14 Point Font.

Dated: May 20, 2022

/s/ Brett A. Mangrum
Brett A. Mangrum

Counsel for Patent Owner-Appellant
Uniloc 2017 LLC